VOLUME 2,NUMBER 3 JULY 2004

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS

EUDOXUS PRESS,LLC

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send three copies of the contribution to the editor in-Chief typed in TEX, LATEX double spaced. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $250,Electronic $220,Print and Electronic $310.Individual:Print $77,Electronic $60,Print &Electronic $110.For any other part of the world add $25 more to the above prices for Print.

Single article PDF file for individual $8.Single issue in PDF form for individual $25.
The journal carries page charges $8 per page of the pdf file of an article,payable upon acceptance of the article within one month and before publication.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

Editorial Board
Associate Editors

tel.515-294-8150
fink@math.iastate.edu
Inequalities,Ordinary Differential Equations

10) Sorin Gal
Department of Mathematics
University of Oradea
Str.Armatei Romane 5
3700 Oradea,Romania
galso@uoradea.ro
Approximation Th.,Fuzzyness,Complex Analysis

11) Jerome A.Goldstein
Department of Mathematical Sciences
The University of Memphis,
Memphis,TN 38152,USA
tel.901-678-2484
jgoldste@memphis.edu
Partial Differential Equations,
Semigroups of Operators

12) Heiner H.Gonska
Department of Mathematics
University of Duisburg
Duisburg,D-47048
Germany
tel.0049-203-379-3542 office
gonska@informatik.uni-duisburg.de
Approximation Th.,Computer Aided Geometric Design

13) Dmitry Khavinson
Department of Mathematical Sciences
University of Arkansas
Fayetteville,AR 72701,USA
tel.(479)575-6331,fax(479)575-8630
dmitry@uark.edu
Potential Th.,Complex Analysis,Holomorphic PDE,Approximation Th.,Function Th.

14) Virginia S.Kiryakova
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms,
Fractional Calculus

15) Hans-Bernd Knoop
Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676

bsendov@bas.bg
Approximation Th.,Geometry of Polynomials,
Image Compression

29) Igor Shevchuk
Faculty of Mathematics and Mechanics
National Taras Shevchenko
University of Kyiv
252017 Kyiv
UKRAINE
shevchuk@univ.kiev.ua
Approximation Theory

30) H.M.Srivastava
Department of Mathematics and Statistics
University of Victoria
Victoria,British Columbia V8W 3P4
Canada
tel.250-721-7455 office,250-477-6960 home,
fax 250-721-8962
harimsri@math.uvic.ca
Real and Complex Analysis,Fractional Calculus and Appl.,
Integral Equations and Transforms,Higher Transcendental
Functions and Appl.,q-Series and q-Polynomials,
Analytic Number Th.

31) Ferenc Szidarovszky
Dept.Systems and Industrial Engineering
The University of Arizona
Engineering Building,111
PO.Box 210020
Tucson,AZ 85721-0020,USA
szidar@sie.arizona.edu
Numerical Methods,Game Th.,Dynamic Systems,
Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources Management

32) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
Approximation Theory

33) Manfred Tasche
Department of Mathematics
University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,

knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

18) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

Image Processing,Harmonic Analysis

34) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

35) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,191-212,
2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Strong And Weak Solutions Of Abstract Cauchy Problems[1]

Chung-Cheng Kuo
Department of Mathematics, Fu-Jen University
Hsin-Chuang, Taipei, Taiwan
e-mail: cckuo@math.fju.edu.tw

Sen-Yen Shaw
Graduate School of Engineering, Lunghwa University of Science and Technology
Gueishan, Taoyuan, Taiwan
e-mail: shaw@math.ncu.edu.tw

Running head: Strong and weak solutions of abstract Cauchy problems

Abstract. Let α be a positive number, C be a bounded linear injection on a Banach space X, and let $A : D(A) \subset X \to X$ be a closed linear operator commuting with C. Under suitable conditions on A (such as, $C^{-1}AC = A$, $\rho(A) \neq \emptyset$, $\overline{D(A)} = X$), we discuss connections among: (i) A being the generator of an α-times integrated C-semigroup on X; (ii) the existence of unique strong solution of $u'(t) = Au(t) + j_{\alpha-1}(t)Cx$, $t > 0$; $u(0) = 0$, for all $x \in D(A)$; (iii) the existence of unique strong solution of $v'(t) = Av(t) + j_\alpha(t)Cx + j_\alpha * Cg(t)$, $t > 0$; $v(0) = 0$, for all $x \in X$; (iv) the existence of unique weak solution of the abstract Cauchy problem: $w'(t) = Aw(t) + j_{\alpha-1}(t)Cx + j_{\alpha-1} * Cg(t)$, $t > 0$; $v(0) = 0$, for all $x \in X$. Here $j_\alpha(t) = t^\alpha/\Gamma(\alpha+1)$ and g is any function in $L^1_{loc}([0,\infty), X)$. Applications to concrete examples are also demonstrated.

2000 *Mathematics Subject Classification*: 47D60, 47D62.
Key words and phrases: α-times integrated C-semigroup, generator, abstract Cauchy problem, strong solution, weak solution.

1 Introduction

Let X be a Banach space with norm $\|\cdot\|$, and let $B(X)$ be the set of all bounded linear operators from X into itself. Consider the abstract Cauchy problem (ACP):

$$\text{(ACP(f,x))} \qquad \begin{cases} u'(t) = Au(t) + f(t), \ t > 0; \\ u(0) = x, \end{cases}$$

[1]Research supported in part by the National Science Council of Taiwan.

where $A : D(A) \subset X \to X$ is a closed linear operator and f is an X-valued function on $[0, \infty)$.

Let X_1 denote the Banach space $D(A)$ equipped with the graph norm $\|x\|_{X_1} = \|x\| + \|Ax\|$ for $x \in D(A)$. A function u is called a *strong solution* of ACP(f, x) if $u \in C^1((0,\infty), X) \cap C([0,\infty), X_1)$ and satisfies ACP(f, x). In the case where A is densely defined, u is called a *weak solution* of ACP(f, x) if u is continuous, and for every $x^* \in D(A^*)$ the function $\langle u(\cdot), x^*\rangle$ is absolutely continuous and satisfies

$$\begin{cases} \frac{d}{dt}\langle u(t), x^*\rangle = \langle u(t), A^*x^*\rangle + \langle f(t), x^*\rangle, \text{a.e.}\, t > 0; \\ u(0) = x. \end{cases}$$

The ACP is closely related to the theory of operator semigroups. Arendt [14, A-II, Theorem 1.1] proved that ACP$(0, x)$ has a unique strong solution for every $x \in D(A)$ if and only if the part A_1 of A in X_1 (i.e. the restriction of A with domain $D(A_1) := \{x \in D(A); Ax \in X_1\}$) generates a C_0-semigroup on X_1. Moreover, these two conditions are also equivalent to that A generates a C_0-semigroup on X, provided that A has nonempty resolvent set $\rho(A)$ [14, A-II, Corollary 1.2]. Ball [2] proved that A has dense domain and ACP(f, x) has a unique weak solution for every $f \in L^1([0,\infty), X)$ and $x \in X$ if and only if A generates a C_0-semigroup on X.

Recently Davies and Pang [3], Miyadera and Tanaka ([16], [17]), deLaubenfels ([4], [5]) have studied C-semigroups (also called regularized semigroups) and their connections with the ACP. A result of deLaubenfels [4, Theorem 4.1] states that if A is the generator of a C-semigroup, then A commutes with C and ACP$(0, x)$ has a unique strong solution for each initial value x in $C(D(A))$. Tanaka and Miyadera [17, Corollary 2.2] then showed that in case $\rho(A) \neq \emptyset$, the converse of the last statement is also true.

The aim of this paper is to prove generalizations of the aforementioned results to α-times ($\alpha > 0$) integrated C-semigroups. This class is a generalization of the class of α-times integrated semigroups, which have been studied in [1], [6], [8], and [15] for $\alpha \in \mathbf{N}$, and [7], [12], and [13] for $\alpha \in \mathbf{R}_+$. This paper serves as a continuation of [9], in which basic properties of α-times integrated C-semigroups as well as characterization of their generators have been discussed.

Let $C \in B(X)$ be an injective operator and α a positive number. A family $S(\cdot) = \{S(t); t \geq 0\}$ in $B(X)$ is called an *α-times integrated C-semigroup* if
(1.1) $S(\cdot)x : [0,\infty) \longrightarrow X$ is continuous for each $x \in X$;
(1.2) $S(0) = 0$, $S(t)C = CS(t)$, and for all $x \in X$, t, $s \geq 0$

$$S(t)S(s)x = \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t + s - r)^{\alpha-1}S(r)Cxdr.$$

An α-times integrated C-semigroup $S(\cdot)$ is said to be *nondegenerate* if
(1.3) $S(t)x = 0$ for all $t > 0$ implies $x = 0$.
Note that for a nondegenerate α-times integrated C-semigroup, the identity $S(0) = 0$ follows automatically from (1.3) and the functional equation in (1.2) and so it is superfluous in condition (1.2) in the nondegenerate case.

The *generator* A of a nondegenerate α-times integrated C-semigroup $S(\cdot)$ is defined as

$$\begin{cases} D(A) := \{x \in X;\ \text{there exists a } y \in X \text{ such that (1.4) is satisfied}\} \\ Ax := y \text{ for } x \in D(A), \end{cases}$$

where

$$\int_0^t S(r)y\,dr = S(t)x - j_\alpha(t)Cx \text{ for all } t \geq 0, \tag{1.4}$$

with $j_\alpha(t) := t^\alpha/\Gamma(\alpha+1)$. Notice that the nondegeneracy implies uniqueness of y in (1.4), so that A is well-defined.

Conditions (1.2) and (1.3) also imply that C is injective. It is known [9] that the generator A is a closed linear operator and has the properties:

$$C^{-1}AC = A; \tag{1.5}$$

$$S(t)A \subset AS(t) \text{ for } t \geq 0 \tag{1.6}$$

and

$$\int_0^t S(s)x\,ds \in D(A) \text{ and } A\int_0^t S(s)x\,ds = S(t)x - j_\alpha(t)Cx \tag{1.7}$$

for all $x \in X$, $t \geq 0$.

A closed linear operator B is called a *subgenerator* of $S(\cdot)$ if it commutes with C and satisfies (1.6) and (1.7) with A therein replaced by B. We have $B \subset C^{-1}BC$, and it can be shown that $C^{-1}BC$ is also a subgenerator. In fact, the generator A is equal to $C^{-1}BC$ for any subgenerator B (cf. [10] for the case $\alpha = n \in \mathbf{N} \cup \{0\}$).

An α-times integrated C-semigroup $S(\cdot)$ is said to be *exponentially bounded* if

(1.8) there are $M \geq 0$ and $w \in \mathbf{R}$ such that $\|S(t)\| \leq Me^{wt}$ for all $t \geq 0$.

In this case, for all $\lambda > w$, $\lambda - A$ is injective, $R(C) \subset R(\lambda - A)$ and

$$(\lambda - A)^{-1}Cx = \lambda^\alpha \int_0^\infty e^{-\lambda t}S(t)x\,dt \text{ for } x \in X. \tag{1.9}$$

Exponentially bounded n-times integrated C-semigroups have been considered in [10] and [11].

In Section 2 we show that, under the assumption that A commutes with C, the problem ACP$(j_{\alpha-1}Cx, 0)$ has a unique strong solution for every $x \in D(A)$ if and only if A_1, the part of A in $X_1 = D(A)$ (i.e. $D(A_1) = D(A^2)$ and $A_1x = Ax$ for $x \in D(A_1)$), is the generator of an α-times integrated C_1-semigroup on X_1, where C_1 is the restriction of C to X_1 (Theorem 2.1). The above condition on A is in general only a necessary condition but not sufficient

condition for A to generate an α-times integrated C-semigroup on X. To seek for a sufficient condition, one may try to enlarge the domain of solvability. It is known [9, Theorem 2.4] that if A generates an exponentially bounded α-times integrated C-semigroup, then for large enough λ, $\lambda - A$ is injective, $R(C) \subset R(\lambda - A)$, and $\mathrm{ACP}(j_{\alpha-1}x, 0)$ has a unique strong solution for each $x \in (\lambda - A)^{-1}C(X)$. Thus the latter condition might be a suitable choice for a sufficient condition. Indeed, this is justified in [9, Theorem 3.2]. An extension (Corollary 2.5) of that result to the case that $\lambda - A$ is not injective will be deduced from Theorem 2.2, which states that if $\mathrm{ACP}(j_\alpha Cx, 0)$ has a unique strong solution for every $x \in X$, then $C^{-1}AC$ generates an α-times integrated C-semigroup on X. Another consequence of Theorem 2.2 asserts that A is the generator of an α-times integrated C-semigroup on X if and only if $C^{-1}AC = A$ and $\mathrm{ACP}(j_\alpha Cx + j_\alpha * Cg, 0)$ has a unique strong solution for every $g \in L^1_{loc}((0,\infty), X)$ and $x \in X$ (see Theorem 2.3).

In Section 3, we discuss connections between the generator of an α-times integrated C-semigroup and weak solutions of the associated ACP. Under the assumption that A has dense domain and $C^{-1}AC = A$, it is shown that A generates an α-times integrated C-semigroup on X if and only if $\mathrm{ACP}(j_{\alpha-1} * Cg + j_{\alpha-1}Cx, 0)$ has a unique weak solution for every $g \in L^1_{loc}((0,\infty), X)$ and $x \in X$, if and only if for every $x \in D(A)$ $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ has a unique strong solution which depends continuously on $x \in D(A)$ (Theorem 3.1).

Some concrete examples of α-times integrated semigroups have been discussed in [6], [7], [8], and [12]. Examples of C-semigroups can be found in [5]. In Section 4, we will apply the present abstract results to two of them.

Finally, we remark that similar results as those in Sections 2 and 3 have been shown ([18] and [19]) to hold for local C-semigroups.

2 Generator and Strong Solutions

This section establishes a precise correspondence between the generator of an α-times integrated C-semigroup and the existence/uniqueness of strong solutions of the corresponding ACP.

It is easy to see that if A is the generator of an α-times integrated C-semigroup, then for every $x \in D(A)$ $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ has the unique strong solution $u(t) = S(t)x$. In general, the converse is not true. However, the following theorem shows that the operator A_1, instead of A, is indeed a generator.

Theorem 2.1. *Let C be a bounded linear injection on X and A be a closed linear operator satisfying*

$$Cx \in D(A) \text{ and } ACx = CAx \text{ for } x \in D(A). \tag{2.1}$$

Then the following statements are equivalent.

(i) $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ *has a unique strong solution for every $x \in D(A)$.*

(ii) *A_1 is the generator of an α-times integrated C_1-semigroup on X_1, where C_1 is the restriction of C to X_1.*

In this case, the solution of $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ *for* $x \in D(A)$ *is given by* $u(\cdot; j_{\alpha-1}Cx, 0) = S_1(\cdot)x$. *Moreover,* $\|S_1(t)\| \leq Me^{wt}$ *for some* M, $w > 0$ *and all* $t \geq 0$ *if and only if* $\|u(t; j_{\alpha-1}Cx, 0)\| = O(e^{wt})\,(t \to \infty)$ *and* $\|u'(t; j_{\alpha-1}Cx, 0)\| = O(e^{wt})\,(t \to \infty)$ *for every* $x \in D(A)$.

Proof. (ii) $\Rightarrow$ (i). Assume that A_1 is the generator of an α-times integrated C_1-semigroup $\{S_1(t); t \geq 0\}$ on X_1. Let $x \in D(A)$ and set $u(t) = S_1(t)x$. Then $u \in C([0,\infty), X_1)$ so that both u and Au are continuous functions. Since A is closed we have $\int_0^t u(s)ds \in D(A)$ and $A\int_0^t u(s)ds = \int_0^t Au(s)ds$. Moreover, $\int_0^t S_1(s)xds \in D(A_1)$ and

$$A\int_0^t u(s)ds = A_1\int_0^t S_1(s)xds = S_1(t)x - j_\alpha(t)C_1x = u(t) - j_\alpha(t)Cx.$$

Consequently, $u(t) = j_\alpha(t)Cx + \int_0^t Au(s)ds$. Hence $u \in C^1([0,\infty), X)$ and $u'(t) = Au(t) + j_{\alpha-1}(t)Cx$. Hence $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ has u as a strong solution. In order to show the uniqueness, assume that u is a solution of $\mathrm{ACP}(0,0)$, we have to show that $u \equiv 0$. Let $v(t) = \int_0^t u(s)ds$. Then the closedness of A implies that $v(t) \in D(A)$ and $Av(t) = \int_0^t Au(s)ds = \int_0^t u'(s)ds = u(t) \in D(A)$. Consequently, $v(t) \in D(A^2) = D(A_1)$ for all $t \geq 0$. Moreover, $v'(t) = u(t) = Av(t)$ and $Av'(t) = Au(t) = u'(t)$ are continuous. Thus $v \in C^1([0,\infty), X_1)$ and $v'(t) = A_1v(t)$, $v(0) = 0$. Since A_1 is assumed to be the generator of an α-times integrated C_1-semigroup $S_1(\cdot)$ on X_1, it follows that $v = S_1(\cdot)0 \equiv 0$ and hence $u \equiv 0$.

(i) $\Rightarrow$ (ii). Assume that for every $x \in D(A)$ there exists a unique solution $u(\cdot; j_{\alpha-1}Cx, 0) \in C^1((0,\infty), X) \cap C([0,\infty), X_1)$ of $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$. For $x \in X_1$, we define $S_1(t)x = u(t; j_{\alpha-1}Cx, 0)$ for $t \geq 0$. By the uniqueness of solution and by (2.1) one can see that $S_1(t)$ is a linear operator on X_1 satisfying $S_1(0) = 0$ and $S_1(\cdot)Cx = CS_1(\cdot)x$. In particular, this implies that $S_1(\cdot)$ commutes with C_1.

Since $S_1(\cdot)x \in C^1((0,\infty), X) \cap C([0,\infty), X_1)$, $S_1(t)x$ is continuous from $[0,\infty)$ into X_1. Now let $C([0,\infty), X_1)$ be the Fréchet space with the quasi-norm $\sum\limits_{k=1}^{\infty} \frac{1}{2^k}\frac{\|v\|_k}{1+\|v\|_k}$, where $\|v\|_k = \max\limits_{t\in[0,k]} \|v(t)\|_{X_1}$ for $k \geq 1$. Consider the linear map $\eta : X_1 \to C([0,\infty), X_1)$ given by $\eta(x) = S_1(\cdot)x$. We will show that η is a closed operator. In fact, let $x_n \to x$ in X_1 and $\eta(x_n) = S_1(\cdot)x_n \to v$ in $C([0,\infty), X_1)$. Then

$$S_1(s)x_n = j_\alpha(s)Cx_n + \int_0^s AS_1(r)x_ndr = j_\alpha(s)Cx_n + A\int_0^s S_1(r)x_ndr.$$

Letting $n \to \infty$ we obtain from the closedness of A and $v \in C([0,\infty), X_1)$ that

$$v(s) = j_\alpha(s)Cx + A\int_0^s v(r)dr = j_\alpha(s)Cx + \int_0^s Av(r)dr$$

for $0 \leq s \leq t$. Uniqueness of solutions of $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ implies that $v(\cdot) = S_1(\cdot)x = \eta(x)$, which shows that η is a closed and hence continuous operator

from X_1 to the space $C([0,\infty), X_1)$. This implies in particular that each $S_1(t)$ is a bounded operator on X_1. Moreover, uniqueness of solutions and the injectivity of C imply that $S_1(\cdot)$ is nondegenerate.

Next, we show that

$$S_1(t)S_1(s)x = \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}S_1(r)C_1 x dr$$

for $x \in X_1$ and $t,\ s \geq 0$. We define

$$v_s(t) = \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}S_1(r)Cxdr$$

for fixed $x \in X$ and $s \geq 0$. Then

$$\begin{aligned}
&Av_s(t)\\
&= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}AS_1(r)Cxdr\\
&= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}\left(\frac{d}{dr}S_1(r)Cx - j_{\alpha-1}(r)C^2x\right)dr.
\end{aligned}$$

For the case $\alpha = 1$ we have

$$\begin{aligned}
Av_s(t) &= \left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(\frac{d}{dr}S_1(r)Cx) - C^2x)dr\\
&= S_1(t+s)Cx - S_1(t)Cx - S_1(s)Cx
\end{aligned}$$

and so

$$v_s'(t) = S_1(t+s)Cx - S_1(t)Cx = Av_s(t) + S_1(s)Cx.$$

If $\alpha > 1$, using integration by parts we have

$$\begin{aligned}
Av_s(t) = &-j_{\alpha-1}(t)S_1(s)Cx - j_{\alpha-1}(s)S_1(t)Cx\\
&+ \frac{1}{\Gamma(\alpha-1)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-2}S_1(r)Cxdr\\
&- \frac{1}{(\Gamma(\alpha))^2}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}r^{\alpha-1}C^2xdr
\end{aligned}$$

and

$$\begin{aligned}
v_s'(t) = &-j_{\alpha-1}(s)S_1(t)Cx\\
&+ \frac{1}{\Gamma(\alpha-1)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-2}S_1(r)Cxdr.
\end{aligned}$$

Since

$$\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}r^{\alpha-1}dr = 0$$

for all t, $s \geq 0$ and $\alpha > 0$ (see [9, Lemma 3.1]), it follows that $v_s'(t) = Av_s(t) + j_{\alpha-1}(t)S_1(s)Cx$. Then uniqueness of solutions implies that $v_s(t) = S_1(t)S_1(s)x$ for all t, $s \geq 0$. Hence

$$\begin{aligned}
&\frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}S_1(r)C_1xdr \\
&= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}S_1(r)Cxdr \\
&= v_s(t) = S_1(t)S_1(s)x
\end{aligned}$$

for all t, $s \geq 0$ and $x \in X_1$.

Now we turn to the case $0 < \alpha < 1$. Hypothesis (i) implies that $v_x(t) = \int_0^t S_1(r)xdr$ is the unique solution of ACP$(j_\alpha Cx, 0)$ for $x \in X_1$. Let $\tilde{S}_1(\cdot)$ be defined by $\tilde{S}_1(t)x := \int_0^t S_1(r)xdr$ for $x \in X_1$ and $t \geq 0$. The previous argument shows that $\tilde{S}_1(\cdot)$ is a nondegenerate $(\alpha+1)$-times integrated C_1-semigroup on X_1. In particular, $A\tilde{S}_1(t)x$ and $\tilde{S}_1(t)x$ are continuous on $[0,\infty)$ for all $x \in X_1$. Since $AS_1(t)x$ is continuous for all $x \in X_1$, the closedness of A implies that $A\tilde{S}_1(t)x = A\int_0^t S_1(r)xdr = \int_0^t AS_1(r)xdr \in C^1([0,\infty), X)$. Hence $\tilde{S}_1(\cdot)x \in C^1([0,\infty), X_1)$. Then differentiation shows that $S_1(\cdot)$ is an α-times integrated C_1-semigroup on X_1.

To see that A_1 is the generator of $S_1(\cdot)$, we first show that

$$AS_1(t)x = S_1(t)Ax \text{ for all } x \in D(A^2) \text{ and } t \geq 0. \tag{2.2}$$

In fact, for a given $x \in D(A^2)$ let $w(t) = j_\alpha(t)Cx + \int_0^t S_1(s)Axds$. Then by the closedness of A and the continuity of the function $AS_1(\cdot)Ax$ and (2.1) we have

$$\begin{aligned}
w'(t) &= j_{\alpha-1}(t)Cx + S_1(t)Ax \\
&= j_{\alpha-1}(t)Cx + j_\alpha(t)CAx + \int_0^t AS_1(s)Axds \\
&= j_{\alpha-1}(t)Cx + j_\alpha(t)ACx + \int_0^t AS_1(s)Axds \\
&= j_{\alpha-1}(t)Cx + Aw(t).
\end{aligned}$$

Since $w(0) = 0$, it follows from uniqueness of solutions that $w(\cdot) \equiv S_1(\cdot)x$. Hence we have

$$AS_1(t)x = Aw(t) = w'(t) - j_{\alpha-1}(t)Cx = S_1(t)Ax,$$

which is (2.2). Now we denote by B the generator of $\{S_1(t); t \geq 0\}$. For $x \in D(A_1) = D(A^2)$ we have by (2.2)

$$\int_0^t S_1(s)A_1xds = \int_0^t AS_1(s)xds = S_1(t)x - j_\alpha(t)Cx$$

for all $t \geq 0$. This shows that $A_1 \subset B$.

In order to show the converse, let $x \in D(B)$. Then

$$A\tilde{S}_1(t)x = S_1(t)x - j_\alpha(t)Cx = \int_0^t S_1(s)Bxds.$$

Differentiating this equation we have $AS_1(t)x = S_1(t)Bx \in D(A)$ for all $t > 0$. Since $AS_1(\cdot)Bx \in C([0,\infty), X)$ and $AS_1(t)Bx = \frac{d}{dt}S_1(t)Bx - j_{\alpha-1}(t)CBx$ for $t > 0$, by the closedness of A we have $\int_0^t S_1(r)Bxdr \in D(A)$ and

$$A\int_0^t S_1(r)Bxdr = \int_0^t AS_1(r)Bxdr = S_1(t)Bx - j_\alpha(t)CBx \in D(A),$$

so that $j_\alpha(t)C_1x = S_1(t)x - \int_0^t S_1(r)Bxdr \in D(A)$ and

$$\begin{aligned} j_\alpha(t)ACx = j_\alpha(t)AC_1x &= A\left(S_1(t)x - \int_0^t S_1(r)Bxdr\right) \\ &= S_1(t)Bx - [S_1(t)Bx - j_\alpha(t)CBx] = j_\alpha(t)CBx. \end{aligned}$$

Since $D(B) \subset X_1 = D(A)$, by (2.1) we have $CAx = ACx = CBx$, so that $Ax = Bx \in X_1 = D(A)$. That means that $x \in D(A_1)$ and $Bx = A_1x$. Consequently, $B \subset A_1$.

Finally, if $\|u(t; j_{\alpha-1}Cx, 0)\| = O(e^{wt})$ and $\|u'(t; j_{\alpha-1}Cx, 0)\| = O(e^{wt})\,(t \to \infty)$ for every $x \in X_1 = D(A)$, then from the equalities $S_1(t)x = u(t; j_{\alpha-1}Cx, 0)$ and

$$\begin{aligned} &\frac{d}{dt}S_1(t)x = AS_1(t)x + j_{\alpha-1}(t)Cx \text{ and} \\ &\|S_1(t)x\|_{X_1} = \|S_1(t)x\| + \|AS_1(t)x\|, \end{aligned} \tag{2.3}$$

we see that $\|S_1(t)x\|_{X_1} = O(e^{wt})\,(t \to \infty)$ for all $x \in X_1$, so that there are M, $w > 0$ such that $\|S_1(t)\| \leq Me^{wt}$ for all $t \geq 0$, by the uniform boundedness principle. The converse is also obvious in view of (2.3). The proof is complete.

The next theorem is a version of Theorem 3.1 of [17] for cases $\alpha > 0$.

Theorem 2.2. *Let C be a bounded linear injection on X and A be a closed linear operator satisfying* (2.1). *Among the assertions below the following implications are valid:* (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii).

(i) *For every $x \in X$, the problem* ACP$(j_\alpha Cx, 0)$ *has a unique strong solution $u(\cdot; j_\alpha Cx, 0) \in C^1([0,\infty), X)$.*

(ii) *The integral equation*

$$v(t) = A\int_0^t v(s)ds + j_\alpha(t)Cx \tag{2.4}$$

has a unique solution $v \in C([0,\infty); X)$ for every $x \in X$.

(iii) $C^{-1}AC$ *is the generator of an* α*-times integrated* C*-semigroup* $S(\cdot)$ *on* X.

Proof. Note that by setting $v(t) = u'(t; j_\alpha Cx, 0)$, one easily sees that (i) and (ii) are equivalent.

(i) $\Rightarrow$ (iii). Assume that for every $x \in X$ there exists a unique strong solution $u(\cdot; j_\alpha Cx, 0)$ of ACP$(j_\alpha Cx, 0)$. Uniqueness of solutions and (2.1) imply that $u(\cdot; j_\alpha C^2 x, 0) = Cu(\cdot; j_\alpha Cx, 0)$. For $x \in X$ and $t, s \geq 0$ we define

$$S(t)x := u'(t; j_\alpha Cx, 0),\ \tilde{S}(t)x := \int_0^t S(r)xdr = u(t; j_\alpha Cx, 0)$$

and

$$v_s(t) := \frac{1}{\Gamma(\alpha+1)} \left(\int_0^{t+s} - \int_0^t - \int_0^s \right) (t+s-r)^\alpha \tilde{S}(r)Cxdr.$$

Clearly both $\tilde{S}(\cdot)x : [0,\infty) \to X$ and $S(\cdot)x : [0,\infty) \to X$ are continuous and nondegenerate. Moreover, the operators $\tilde{S}(t)$ and $S(t)$ are linear and commute with C. Using integration by parts we have

$$\begin{aligned}
Av_s(t) &= \frac{1}{\Gamma(\alpha+1)} \left(\int_0^{t+s} - \int_0^t - \int_0^s \right) (t+s-r)^\alpha A\tilde{S}(r)Cxdr \\
&= \frac{1}{\Gamma(\alpha+1)} \left(\int_0^{t+s} - \int_0^t - \int_0^s \right) (t+s-r)^\alpha \left(S(r)Cx - j_\alpha(r)C^2x \right) dr \\
&= -j_\alpha(t)C\tilde{S}(r)Cx - j_\alpha(s)\tilde{S}(t)Cx \\
&\quad + \frac{1}{\Gamma(\alpha)} \left(\int_0^{t+s} - \int_0^t - \int_0^s \right) (t+s-r)^{\alpha-1} \tilde{S}(r)Cxdr \\
&\quad - \frac{1}{(\Gamma(\alpha+1))^2} \left(\int_0^{t+s} - \int_0^t - \int_0^s \right) (t+s-r)^\alpha r^\alpha C^2xdr
\end{aligned}$$

and

$$v_s'(t) = -j_\alpha(s)\tilde{S}(t)Cx + \frac{1}{\Gamma(\alpha)} \left(\int_0^{t+s} - \int_0^t - \int_0^s \right) (t+s-r)^{\alpha-1} \tilde{S}(r)Cxdr.$$

Since $\left(\int_0^{t+s} - \int_0^t - \int_0^s \right) (t+s-r)^\alpha r^\alpha dr = 0$ for all $t, s \geq 0$ and $\alpha > 0$ (see [9]), it follows that $v_s'(t) = Av_s(t) + j_\alpha(t)C\tilde{S}(s)x$ for all $t \geq 0$. Then the uniqueness of solution implies that $v_s(t) = \tilde{S}(t)\tilde{S}(s)x$ for all $t, s \geq 0$.

Now let $C([0,\infty), X)$ be the Fréchet space with quasi-norm $\sum\limits_{k=1}^{\infty} \frac{1}{2^k} \frac{\|v\|_k}{1+\|v\|_k}$, where $\|v\|_k = \max\{\|v(t)\|_X; t \in [0,k]\}$ for $k \geq 1$. Consider the linear map $\eta : X \to C([0,\infty), X)$ given by $\eta(x) = \tilde{S}(\cdot)x$. We will show that η is a closed operator. In fact, let $x_n \to x$ in X and $\eta(x_n) = \tilde{S}(\cdot)x_n \to v$ in $C([0,\infty), X)$. Then for each $k \in \mathbf{N}$, $\tilde{S}(t)x_n \to v(t)$ and

$$A\tilde{S}(t)x_n = S(t)x_n - j_\alpha(t)Cx_n \to Av(t)$$

uniformly on $t \in [0,k]$, so that $v(0)=0$ and $S(t)x_n \to Av(t)+j_\alpha(t)Cx$ uniformly on $[0,k]$. Hence v is differentiable on $[0,k)$ for each $k>0$ and $\frac{d}{dt}v(t)=Av(t)+j_\alpha(t)Cx$. By uniqueness of solutions, $v(\cdot)=\tilde{S}(\cdot)x=\eta(x)$. This shows that η is closed and hence is a continuous operator from X to $C([0,\infty);X)$. In particular, both $\tilde{S}(t)$ and $S(t)$ belong to $B(X)$ for each $t\geq 0$.

Having shown that $\tilde{S}(\cdot)$ is an $(\alpha+1)$-times integrated C-semigroup, we now prove that $S(\cdot)$ is an α-times integrated C-semigroup. Indeed, differentiation with respect to s yields

$$\begin{aligned}
&\tilde{S}(t)S(s)x=\tilde{S}(t)\frac{d}{ds}\tilde{S}(s)x=\frac{d}{ds}\tilde{S}(t)\tilde{S}(s)x\\
&=\frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s}-\int_0^t-\int_0^s\right)(t+s-r)^{\alpha-1}\tilde{S}(r)Cxdr-j_\alpha(t)\tilde{S}(s)Cx.\\
&=j_\alpha(s)\tilde{S}(t)Cx+\frac{1}{\Gamma(\alpha+1)}\left(\int_0^{t+s}-\int_0^t-\int_0^s\right)(t+s-r)^{\alpha}S(r)Cxdr.
\end{aligned}$$

Next, we take derivatives with respect to t and get

$$S(t)S(s)x=\frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s}-\int_0^t-\int_0^s\right)(t+s-r)^{\alpha-1}S(r)Cxdr.$$

Moreover, $S(0)x=A\tilde{S}(0)x=0$ for all $x\in X$. Consequently, $S(\cdot)$ is an α-times integrated C-semigroup.

Finally, we show that $C^{-1}AC$ is the generator of $\{S(t);t\geq 0\}$. Let B be the generator of $S(\cdot)$ and let $x\in D(B)$. Then

$$A\tilde{S}(t)x=S(t)x-j_\alpha(t)Cx=\int_0^t S(r)Bxdr=\tilde{S}(t)Bx \text{ for } t\geq 0.$$

By the closedness of A and the fact $\tilde{S}(\cdot)Bx\in C^1([0,\infty),X)\cap C([0,\infty),X_1)$ we have $\int_0^t \tilde{S}(r)Bxdr\in D(A)$ and

$$\begin{aligned}
A\int_0^t \tilde{S}(r)Bxdr&=\int_0^t A\tilde{S}(r)Bxdr\\
&=\int_0^t\left(\frac{d}{dr}\tilde{S}(r)Bx-j_\alpha(r)CBx\right)dr\\
&=\tilde{S}(t)Bx-j_{\alpha+1}(t)CBx.
\end{aligned}$$

Since

$$j_{\alpha+1}(t)Cx=-\int_0^t\tilde{S}(r)Bxdr+\tilde{S}(t)x\in D(A),$$

it follows that $Cx\in D(A)$ and

$$\begin{aligned}
j_{\alpha+1}(t)ACx&=-A\int_0^t\tilde{S}(r)Bxdr+A\tilde{S}(t)x\\
&=-\tilde{S}(t)Bx+j_{\alpha+1}(t)CBx+\tilde{S}(t)Bx=j_{\alpha+1}(t)CBx.
\end{aligned}$$

Consequently, $x \in D(C^{-1}AC)$ and $C^{-1}ACx = Bx$.

Conversely, let $x \in D(C^{-1}AC)$ and consider

$$w(t) = \int_0^t \tilde{S}(s)C^{-1}ACxds + j_{\alpha+1}(t)Cx.$$

Since

$$\begin{aligned} S(s)C^{-1}ACx &= A\tilde{S}(s)C^{-1}ACx + j_\alpha(t)C(C^{-1}AC)x \\ &= A\left(\tilde{S}(s)C^{-1}ACx + j_\alpha(t)Cx\right), \end{aligned}$$

taking derivatives and integrals yields

$$\begin{aligned} w'(t) - j_\alpha(t)Cx &= \tilde{S}(t)C^{-1}ACx \\ &= \int_0^t A\left(\tilde{S}(s)C^{-1}ACx + j_\alpha(s)Cx\right)ds \\ &= Aw(t). \end{aligned}$$

Uniqueness of solutions shows that $w(t) = \tilde{S}(t)x$ for $t \in [0,\infty)$. Therefore

$$\begin{aligned} S(t)x = w'(t) &= \tilde{S}(t)C^{-1}ACx + j_\alpha(t)Cx \\ &= \int_0^t S(r)C^{-1}ACxdr + j_\alpha(t)Cx \end{aligned}$$

for all $t \geq 0$. This means that $x \in D(B)$ and $Bx = C^{-1}ACx$. It follows that $B = C^{-1}AC$.

Theorem 2.3. *Let C be a bounded linear injection on X and let A be a closed linear operator. Then the following statements are equivalent.*

(i) *A is the generator of an α-times integrated C-semigroup $S(\cdot)$ on X.*

(ii) *$C^{-1}AC = A$ and the problem $\mathrm{ACP}(j_\alpha Cx + j_\alpha * Cg, 0)$ has a unique strong solution $u(\cdot; j_\alpha Cx + j_\alpha * Cg, 0)$ for every $g \in L^1_{loc}([0,\infty), X)$ and $x \in X$.*

(ii') *$C^{-1}AC = A$, and the integral equation*

$$v(t) = A\int_0^t v(s)ds + j_\alpha(t)Cx + C(j_\alpha * g)(t) \tag{2.5}$$

has a unique solution $v \in C([0,\infty); X)$ for every $g \in L^1_{loc}([0,\infty), X)$ and $x \in X$.

(iii) *$C^{-1}AC = A$, and the problem $\mathrm{ACP}(j_\alpha Cx, 0)$ has a unique strong solution $u(\cdot; j_\alpha Cx, 0) \in C^1([0,\infty), X)$ for every $x \in X$.*

(iii') *$C^{-1}AC = A$, and the integral equation (2.4) has a unique solution $v \in C([0,\infty); X)$ for every $x \in X$.*

The solution $u(\cdot; j_\alpha Cx + j_\alpha * Cg, 0)$ is given by

$$u(t; j_\alpha Cx + j_\alpha * Cg, 0) = \int_0^t S(s)xds + \int_0^t \int_0^s S(s-r)g(r)drds.$$

Moreover, $S(\cdot)$ is exponentially bounded if and only if $\|u(t; j_\alpha Cx, 0)\| = O(e^{wt})$ and $\|u'(t; j_\alpha Cx, 0)\| = O(e^{wt})$ $(t \to \infty)$ for some $w > 0$ and all $x \in X$.

Proof. Note that by setting $v(t) = u'(t; j_\alpha Cx + j_\alpha * Cg, 0)$, one sees that (ii) and (ii') are equivalent. "(ii) $\Rightarrow$(iii)" is obvious. Thus, in view of Theorem 2.2, it remains to show "(i) $\Rightarrow$ (ii)". For this, in view of (1.4) and (1.7), we need only to show that $u(t) = \int_0^t \int_0^s S(s-r)g(r)drds$ satisfies ACP$(j_\alpha * Cg, 0)$. Using (1.7) and the closedness of A we have

$$\begin{aligned} Au(t) &= A\int_0^t \int_0^s S(s-r)g(r)drds = A\int_0^t \int_r^t S(s-r)g(r)dsdr \\ &= \int_0^t A\int_0^{t-r} S(s)g(r)dsdr = \int_0^t (S(t-r)g(r) - j_\alpha(t-r)Cg(r))\,dr \\ &= u'(t) - j_\alpha * Cg(t) \end{aligned}$$

Uniqueness of solutions for ACP$(j_\alpha Cx + j_\alpha * Cg, 0)$ follows from the unique existence of strong solution of ACP$(0,0)$ and hence the proof is complete.

Applying Theorems 2.1 and 2.3 we prove the following result.

Corollary 2.4. *Let C be a bounded linear injection on X. The statements below are related as follows:* (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) $\Leftrightarrow$ (iv).

(i) *A is the generator of an α-times integrated C-semigroup $S(\cdot)$ on X.*

(ii) *$C^{-1}AC = A$ and the problem* ACP$(j_\alpha Cx + j_\alpha * Cg, 0)$ *has a unique strong solution $u(\cdot; j_\alpha Cx + j_\alpha * Cg, 0)$ for every $g \in L^1_{loc}([0,\infty), X)$ and $x \in X$.*

(iii) *A is a closed linear operator satisfying* (2.1)*, and A_1 is the generator of an α-times integrated C_1-semigroup $S_1(\cdot)$ on X_1.*

(iv) *A is a closed linear operator satisfying* (2.1)*, and for every $x \in D(A)$ there exists a unique strong solution $u(\cdot; j_{\alpha-1}Cx, 0)$ of* ACP$(j_{\alpha-1}Cx, 0)$.

In case A has nonempty resolvent set, all the above statements are equivalent. Moreover, $S_1(\cdot)$ is the restriction of $S(\cdot)$ to X_1, and $u(\cdot; j_{\alpha-1}Cx, 0) = S(t)x$ for $x \in D(A)$.

Proof. "(i) $\Leftrightarrow$ (ii)" follows from Theorem 2.3, and "(iii) $\Leftrightarrow$ (iv)" follows from Theorem 2.1.

(i) $\Rightarrow$ (iii). Suppose A generates an α-times integrated C-semigroup $S(\cdot)$. (1.5) implies (2.1). Let $S_1(t) := S(t)|X_1$ for $t \geq 0$. It is easy to see that $S_1(\cdot)$ is an α-times integrated C_1-semigroup on X_1, with $C_1 = C|X_1$. To show that its generator B is equal to A_1, first let $x \in D(A_1) = D(A^2)$. Then we have

$$\int_0^t S_1(r)A_1xdr = \int_0^t S(r)Axdr = S(t)x - j_\alpha(t)Cx = S_1(t)x - j_\alpha(t)C_1x,$$

which shows that $x \in D(B)$ and $Bx = A_1x$. Hence $A_1 \subset B$. Conversely, if $x \in D(B)$, then

$$S(t)x - j_\alpha(t)Cx = S_1(t)x - j_\alpha(t)C_1x = \int_0^t S_1(r)Bxdr = \int_0^t S(r)Bxdr,$$

so that $x \in D(A)$ and $Ax = Bx \in X_1 = D(A)$. Hence $D(B) \subset D(A^2) = D(A_1)$.

Next, we show "(iii) $\Rightarrow$ (i)" under the assumption $\rho(A) \neq \emptyset$. Suppose A_1 is the generator of an α-times integrated C_1-semigroup $S_1(\cdot)$ on X_1. Let $\lambda \in \rho(A)$, and define $S(\cdot) = (\lambda - A)S_1(\cdot)(\lambda - A)^{-1}$. Since $(\lambda - A)^{-1}$ is an isomorphism from X onto X_1, the strong continuity of $S_1(\cdot)$ on $[0, \infty)$ in X_1 implies the strong continuity of $S(\cdot)$ on $[0, \infty)$ in X. Clearly, $S(0) = 0$, $S(\cdot)$ commutes with $C := (\lambda - A)C_1(\lambda - A)^{-1}$, and

$$\begin{aligned}
&S(t)S(s)x = (\lambda - A)S_1(t)(\lambda - A)^{-1}(\lambda - A)S_1(s)(\lambda - A)^{-1}x \\
&= (\lambda - A)S_1(t)S_1(s)(\lambda - A)^{-1}x \\
&= \frac{1}{\Gamma(\alpha)}(\lambda - A)\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t + s - r)^{\alpha-1}S_1(r)C_1(\lambda - A)^{-1}xdr \\
&= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t + s - r)^{\alpha-1}(\lambda - A)S_1(r)(\lambda - A)^{-1} \\
&\qquad\qquad (\lambda - A)C_1(\lambda - A)^{-1}xdr \\
&= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t + s - r)^{\alpha-1}S(r)Cxdr
\end{aligned}$$

for all $t, s \geq 0$ and $x \in X$. Hence $S(\cdot)$ is an α-times integrated C-semigroup on X. Let G be its generator. If $x \in D(A)$, then $(\lambda - A)^{-1}x \in D(A^2) = D(A_1)$, so that

$$\begin{aligned}
S(t)x - j_\alpha(t)Cx &= (\lambda - A)\left(S_1(t)(\lambda - A)^{-1}x - j_\alpha(t)C_1(\lambda - A)^{-1}x\right) \\
&= (\lambda - A)\int_0^t S_1(r)A_1(\lambda - A)^{-1}xdr \\
&= \int_0^t (\lambda - A)S_1(r)A_1(\lambda - A)^{-1}xdr \\
&= \int_0^t S(r)(\lambda - A)A(\lambda - A)^{-1}xdr \\
&= \int_0^t S(r)Axdr,
\end{aligned}$$

which means that $x \in D(G)$ and $Gx = Ax$. Hence $A \subset G$. Conversely, if $x \in D(G)$, then

$$\begin{aligned}
\int_0^t S_1(r)(\lambda - A)^{-1}Gxdr &= \int_0^t (\lambda - A)^{-1}S(r)Gxdr \\
&= (\lambda - A)^{-1}\int_0^t S(r)Gxdr \\
&= (\lambda - A)^{-1}\left(S(t)x - j_\alpha(t)Cx\right) \\
&= S_1(t)(\lambda - A)^{-1}x - j_\alpha(t)C_1(\lambda - A)^{-1}x.
\end{aligned}$$

That means $(\lambda - A)^{-1}x \in D(A_1) = D(A^2)$ and $A_1(\lambda - A)^{-1}x = (\lambda - A)^{-1}Gx$. Therefore $x \in D(A)$ and $Gx = (\lambda - A)A_1(\lambda - A)^{-1}x = Ax$ for $x \in D(G)$, and we have shown that A is the generator of $S(\cdot)$.

Remarks. In the case where $\alpha = 0$ and $C = I$, this corollary reduces to Corollary 1.2 of [14, A-II]. That (i) implies (iv) for the case $\alpha = 0$ was proved by deLaubenfels [4, Theorem 4.1]. The equivalence of (i) and (iv) (and hence (iii)) in the case that $\alpha = 0$ and $\rho(A) \neq \emptyset$ was proved by Tanaka and Miyadera [17, Corollary 2.2]. It will be seen in Theorem 3.1 that under the condition: $\overline{D(A)} = X$ and $C^{-1}AC = A$, (i) is equivalent to (iv) together with the additional assumption of continuous dependency of solutions on initial values.

Finally, as an application of Theorem 2.2, we deduce the following corollary which improves Theorem 3.2 of [9] by removing the injectivity assumption on $\lambda - A$.

Corollary 2.5. *Let A be a closed linear operator satisfying* (2.1)*. Assume that for some $\lambda \in \mathbf{C}$, $R(C) \subset R(\lambda - A)$, and* $\mathrm{ACP}(j_{\alpha-1}x, 0)$ *has a unique strong solution in $C^1([0,\infty), X)$ for each $x \in D(A)$ such that $(\lambda - A)x \in R(C)$. Then there exists an α-times integrated C-semigroup $S(\cdot)$ on X with generator $C^{-1}AC$.*

Proof. Due to Theorem 2.2, we only have to show that for any given $x \in X$ the integral equation

$$v(t) = A\int_0^t v(s)ds + j_\alpha(t)Cx \tag{2.6}$$

has a unique solution $v \in C([0,\infty); X)$. By assumption, there is a $y \in D(A)$ such that $(\lambda - A)y = Cx$, and $\mathrm{ACP}(j_{\alpha-1}y, 0)$ has a unique strong solution $u(\cdot; y)$. Thus $u(\cdot; y) \in C([0,\infty), [D(A)])$ and $u'(t; y) = Au(t; y) + j_{\alpha-1}(t)y$ for $t > 0$ and $u(0; y) = 0$. The closedness of A and the continuity of $Au(\cdot; y)$ imply that $\int_0^t u(s; y)ds \in D(A)$ and

$$A\int_0^t u(s;y)ds = \int_0^t Au(s;y)ds = u(t;y) - j_\alpha(t)y \in D(A),$$

so that

$$\begin{aligned}(\lambda - A)u(t;y) &= (\lambda - A)A\int_0^t u(s;y)ds + (\lambda - A)j_\alpha(t)y \\ &= A\int_0^t (\lambda - A)u(s;y)ds + j_\alpha(t)Cx\end{aligned}$$

for all $t \geq 0$. That is, $v(t) := (\lambda - A)u(t; y)$ is a solution of (2.6). Uniqueness of solutions for (2.6) follows from uniqueness of solutions for $\mathrm{ACP}(0, 0)$.

3 Generator and Weak Solutions

In this section we assume that A has dense domain and $C^{-1}AC = A$. For $\alpha > 0$ and $g \in L^1_{loc}([0,\infty), X)$ the integral $j_{\alpha-1} * g(t) := \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} g(s)ds$ exists for almost all $t \geq 0$ and $j_{\alpha-1} * g \in L^1_{loc}([0,\infty), X)$; it belongs to $C([0,\infty), X)$ if g does. Moreover, $g \to 0$ in $L^1_{loc}([0,\infty), X)$ implies $j_{\alpha-1} * g \to 0$ in $L^1_{loc}([0,\infty), X)$. A characterization of generators of α-times integrated C-semigroups in terms of unique existence of weak solutions of $\mathrm{ACP}(j_{\alpha-1} * Cg + j_{\alpha-1}Cx, 0)$ and of strong solutions of $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ for all $x \in X$, and strong solutions of $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ for all $x \in D(A)$ are given by the following theorem. Note that the equivalence of (i) and (ii) for the case that $\alpha = 0$ and $C = I$ can be found in [2].

Theorem 3.1. *Let C be a bounded linear injection on X, and A be a densely defined closed linear operator such that $C^{-1}AC = A$. Then the following statements are equivalent:*

(i) *A generates an α-times integrated C-semigroup $S(\cdot)$ on X.*

(ii) *For every $g \in L^1_{loc}([0,\infty), X)$ and $x \in X$ there exists a unique weak solution u of $\mathrm{ACP}(j_{\alpha-1} * Cg + j_{\alpha-1}Cx, 0)$, i.e., there exists $u \in C([0,\infty), X)$ satisfying satisfies*

$$(3.1) \quad \begin{cases} \frac{d}{dt}\langle u(t), x^* \rangle = \langle u(t), A^*x^* \rangle + \langle C(j_{\alpha-1} * g + j_{\alpha-1}(t)x), x^* \rangle \ a.e.\ t \geq 0, \\ u(0) = 0 \end{cases}$$

for all $x^* \in D(A^*)$.

(iii) *For every $x \in X$, $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ has a unique weak solution w.*

(iv) *For every $x \in D(A)$ $\mathrm{ACP}(j_{\alpha-1}Cx, 0)$ has a unique strong solution $u(\cdot; j_{\alpha-1}Cx, 0)$ which depends continuously on x, i.e. if $\{x_n\}$ is a Cauchy sequence in $D(A)$, then $\{u(\cdot; j_{\alpha-1}Cx_n, 0)\}$ is uniformly Cauchy on compact subsets of $[0,\infty)$.*

Moreover, the unique weak solution u of $\mathrm{ACP}(j_{\alpha-1} * Cg + j_{\alpha-1}Cx, 0)$ for $x \in X$ is given by

$$u(t) = S(t)x + \int_0^t S(t-s)g(s)ds, \ t \geq 0.$$

Proof. (i) $\Rightarrow$ (ii). Let A generate the α-times integrated C-semigroup $\{S(t) : t \geq 0\}$. By (1.7) we have for all $x \in X$ and $x^* \in D(A^*)$

$$\begin{aligned} &\langle \frac{1}{h}\big(S(t+h)x - S(t)x\big), x^* \rangle \\ &= \langle \frac{1}{h} \int_t^{t+h} S(s)xds, A^*x^* \rangle + \frac{(t+h)^\alpha - t^\alpha}{h} \frac{1}{\Gamma(\alpha+1)} \langle Cx, x^* \rangle \\ &\to \langle S(t)x, A^*x^* \rangle + \langle j_{\alpha-1}(t)Cx, x^* \rangle \text{ as } h \to 0, \end{aligned}$$

so that $\frac{d}{dt}\langle S(t)x, x^* \rangle = \langle S(t)x, A^*x^* \rangle + \langle j_{\alpha-1}(t)Cx, x^* \rangle$ for all $t > 0$, $x \in X$ and $x^* \in D(A^*)$. Suppose that $g \in C([0,\infty), X)$ and $x \in X$, and let

$$u(t) = S(t)x + \int_0^t S(t-s)g(s)ds.$$

Then $u \in C([0,\infty),X), u(0)=0$, and

$$\langle u(t),x^*\rangle = \langle S(t)x,x^*\rangle + \int_0^t \langle S(t-s)g(s),x^*\rangle ds.$$

It follows that

$$\begin{aligned}
\frac{d}{dt}\langle u(t),x^*\rangle &= \langle S(t)x,A^*x^*\rangle + \langle j_{\alpha-1}(t)Cx,x^*\rangle + \frac{d}{dt}\int_0^t \langle S(t-s)g(s),x^*\rangle ds \\
&= \langle S(t)x,A^*x^*\rangle + \langle j_{\alpha-1}(t)Cx,x^*\rangle + \int_0^t \langle S(t-s)g(s),A^*x^*\rangle ds \\
&\quad + \int_0^t \langle j_{\alpha-1}(t-s)Cg(s),x^*\rangle ds \\
&= \langle u(t),A^*x^*\rangle + \langle C(j_{\alpha-1}*g(t)+j_{\alpha-1}(t)x),x^*\rangle
\end{aligned}$$

for all $t>0$. If $g \in L^1_{loc}([0,\infty),X)$, we choose $g_m \in C([0,\infty),X)$ such that $g_m \to g$ in $L^1_{loc}([0,\infty),X)$, and define

$$u_m(t) = S(t)x + \int_0^t S(t-s)g_m(s)ds,$$

and

$$u(t) = S(t)x + \int_0^t S(t-s)g(s)ds.$$

Then

$$\|u_m(t)-u(t)\| \le \int_0^r \sup_{\sigma\in[0,r]} \|S(\sigma)\| \, \|g_m(s)-g(s)\| \, ds$$

for all $0 \le t \le r$, so that $u_m(\cdot) \to u(\cdot)$ uniformly on compact subsets of $[0,\infty)$. For each $x^* \in D(A^*)$, $\frac{d}{dt}\langle u_m(t),x^*\rangle = \langle u_m(t),A^*x^*\rangle + \langle Cj_{\alpha-1}*g_m(t),x^*\rangle + \langle j_{\alpha-1}(t)Cx,x^*\rangle$, so that

$$\begin{aligned}
\langle u_m(t),x^*\rangle &= \int_0^t \langle u_m(s),A^*x^*\rangle ds + \int_0^t \langle Cj_{\alpha-1}*g_m(s),x^*\rangle ds \\
&\quad + \int_0^t \langle j_{\alpha-1}(s)Cx,x^*\rangle ds.
\end{aligned}$$

Letting $m\to\infty$, we have

$$\begin{aligned}
\langle u(t),x^*\rangle &= \int_0^t \langle u(s),A^*x^*\rangle ds + \int_0^t \langle Cj_{\alpha-1}*g(s),x^*\rangle ds \\
&\quad + \int_0^t \langle j_{\alpha-1}(s)Cx,x^*\rangle ds,
\end{aligned}$$

so that $\frac{d}{dt}\langle u(t),x^*\rangle = \langle u(t),A^*x^*\rangle + \langle Cj_{\alpha-1}*g(t),x^*\rangle + \langle j_{\alpha-1}(t)Cx,x^*\rangle$ for almost all $t \ge 0$ and all $x^* \in D(A^*)$. Hence $u(t) \in C([0,\infty),X)$ satisfies (3.1).

To prove the uniqueness, let $v(t)$ be another solution of (3.1) and let $w(t) = u(t) - v(t)$. Then $\langle w(t), x^*\rangle = \langle \int_0^t w(s)ds, A^*x^*\rangle$ for all $x^* \in D(A^*), t \geq 0$. This implies that $y(t) = \int_0^t w(s)ds$ belongs to $D(A)$ and $Ay(t) = w(t) = y'(t)$. Since $y(0) = 0$ we must have $y \equiv 0$ and hence $u(\cdot) \equiv v(\cdot)$.

(iii) $\Rightarrow$ (i). For any $x \in X$ let $w(\cdot; j_{\alpha-1}Cx, 0)$ be the unique weak solution of the problem ACP$(j_{\alpha-1}Cx, 0)$ (i.e. the unique solution of (3.1) with $g \equiv 0$) and define for $t \geq 0$ the map $S(t) : X \to X$ by $S(t)x = w(t; j_{\alpha-1}Cx, 0)$ $(x \in X)$. Then $S(\cdot)x \in C([0,\infty), X)$ for all $x \in X$. Also, by the uniqueness of solution one easily infers that $S(t)$ is linear and nondegenerate, $S(0) = 0$, and $S(\cdot)$ commutes with C so that $w(t; C^2x) = S(t)Cx = CS(t)x = Cw(t; j_{\alpha-1}Cx, 0)$.

In order to show that for each $t \geq 0$ the operator $S(t)$ is bounded, we consider the linear map $\eta : X \to C([0,\infty), X)$ given by $\eta(x) = S(\cdot)x$. Then the operator η is continuous. Indeed, by the closed graph theorem it suffices to show that η is closed. To this end, let x_m be a sequence such that $x_m \to x$ in X and $\eta(x_m) = S(\cdot)x_m \to u(\cdot)$ in $C([0,\infty), X)$ as $m \to \infty$. Then, from the equality

$$\begin{aligned}\int_0^t \langle S(s)x_m, A^*x^*\rangle ds &= \int_0^t \frac{d}{ds}\langle S(s)x_m, x^*\rangle ds - \int_0^t \langle j_{\alpha-1}(s)Cx_m, x^*\rangle ds \\ &= \langle S(t)x_m, x^*\rangle - \langle j_\alpha(t)Cx_m, x^*\rangle\end{aligned}$$

it follows that as $m \to \infty$

$$\langle u(t), x^*\rangle = \langle j_\alpha(t)Cx, x^*\rangle + \int_0^t \langle u(s), A^*x^*\rangle ds \tag{3.2}$$

for $t \geq 0$ and $x^* \in D(A^*)$. That is, u is a weak solution of ACP$(j_{\alpha-1}Cx, 0)$. Therefore, from uniqueness of weak solutions it follows that $u(\cdot) = S(\cdot)x = \eta(x)$. Hence η is closed.

To show that the family $\{S(t); t \geq 0\}$ satisfies (1.2), we define for fixed $x \in X$ and $s \geq 0$

$$v_s(t) = \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}S(r)Cxdr.$$

Then

$$\begin{aligned}\langle v_s(t), A^*x^*\rangle &= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}\langle S(r)Cx, A^*x^*\rangle dr \\ &= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1} \\ &\qquad \left(\frac{d}{dr}\langle S(r)Cx, x^*\rangle - \langle j_{\alpha-1}(r)C^2x, x^*\rangle\right)dr \\ &= \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}\frac{d}{dr}\langle S(r)Cx, x^*\rangle dr \\ &\qquad - \langle \frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}j_{\alpha-1}(r)C^2xdr, x^*\rangle.\end{aligned}$$

For $\alpha = 1$ we have

$$\langle v_s(t), A^*x^*\rangle = \langle S(t+s)Cx - S(t)Cx - S(s)Cx, x^*\rangle$$

and

$$\frac{d}{dt}\langle v_s(t), x^*\rangle = \langle S(t+s)Cx - S(t)Cx, x^*\rangle.$$

For $\alpha > 1$ we have

$$\begin{aligned}\langle v_s(t), A^*x^*\rangle = &-j_{\alpha-1}(t)\langle S(s)Cx, x^*\rangle - j_{\alpha-1}(s)\langle S(t)Cx, x^*\rangle \\ &+ \frac{1}{\Gamma(\alpha-1)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-2}\langle S(r)Cx, x^*\rangle dr \\ &- \frac{1}{\Gamma(\alpha)}\langle\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}j_{\alpha-1}(r)C^2x dr, x^*\rangle\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}\langle v_s(t), x^*\rangle = &-j_{\alpha-1}(s)\langle S(t)Cx, x^*\rangle \\ &+ \frac{1}{\Gamma(\alpha-1)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-2}\langle S(r)Cx, x^*\rangle dr.\end{aligned}$$

Using the fact that

$$\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}r^{\alpha-1}dr = 0 \text{ for all } t,\ s \geq 0 \text{ and } \alpha > 0,$$

one obtains

$$\frac{d}{dt}\langle v_s(t), x^*\rangle = \langle v_s(t), A^*x^*\rangle + \langle j_{\alpha-1}(t)CS(s)x, x^*\rangle$$

for all $t \geq 0$ and $v_s(0) = 0$. Then uniqueness of solutions implies that $v_s(t) = S(t)S(s)x$ and hence

$$\frac{1}{\Gamma(\alpha)}\left(\int_0^{t+s} - \int_0^t - \int_0^s\right)(t+s-r)^{\alpha-1}S(r)Cx dr = v_s(t) = S(t)S(s)x$$

for all $t, s \geq 0$ and $x \in X$.

Now we turn to the case $0 < \alpha < 1$. Hypothesis (iii) implies that $\tilde{S}(t)x = \int_0^t S(r)x dr$ is the unique weak solution of ACP$(j_\alpha Cx, 0)$ for $x \in X$ and $\tilde{S}(\cdot)x \in C^1([0,\infty), X)$. Then an easy computation shows that $S(\cdot)$ is an α-times integrated C-semigroup.

Consequently, $S(\dot{})$ is a nondegenerate α-times integrated C-semigroup on X. Finally we show that A is the generator of $S(\dot{})$. Let B be the generator of $S(\cdot)$ and let $x \in D(B)$. For any $x^* \in D(A^*)$, we have

$$\langle S(t)x, A^*x^*\rangle = \frac{d}{dt}\langle S(t)x, x^*\rangle - \langle j_{\alpha-1}(t)Cx, x^*\rangle = \langle S(t)Bx, x^*\rangle.$$

It follows that $S(t)x \in D(A)$ and $AS(t)x = S(t)Bx$. Since

$$\begin{aligned}\langle \int_0^t S(r)Bxdr, A^*x^* \rangle &= \int_0^t \left(\frac{d}{dr}\langle S(r)Bx, x^*\rangle - \langle j_{\alpha-1}(r)CBx, x^*\rangle \right) dr \\ &= \langle S(t)Bx, x^*\rangle - \langle j_\alpha(t)CBx, x^*\rangle\end{aligned}$$

for all $x^* \in D(A^*)$, we have $\int_0^t S(r)Bxdr \in D(A)$ and $A\int_0^t S(r)Bxdr = S(t)Bx$ $-j_\alpha(t)CBx$, which implies that $-j_\alpha(t)Cx = \int_0^t S(r)Bxdr - S(t)x \in D(A)$ and

$$\begin{aligned}-j_\alpha(t)ACx &= A\int_0^t S(r)Bxdr - AS(t)x \\ &= S(t)Bx - j_\alpha(t)CBx - S(t)Bx \\ &= -j_\alpha(t)CBx\end{aligned}$$

for $t > 0$. This shows that $B \subset C^{-1}AC = A$. To prove the opposite inclusion, let $x \in D(A)$. Using $B \subset A$ we see that for each $t \geq 0$, the integrals $\int_0^t S(s)xds$ and $\int_0^t S(s)Axds$ belong to $D(A)$ and

$$S(t)x = j_\alpha(t)Cx + A\int_0^t S(s)xds, \tag{3.3}$$

$$S(t)Ax = j_\alpha(t)CAx + A\int_0^t S(s)Axds. \tag{3.4}$$

Consider the function

$$v(t) = \int_0^t S(s)Axds - A\int_0^t S(s)xds.$$

It follows from (3.3) that $v(\cdot) \in C([0,\infty), X)$ and $v(0) = 0$. Now, using (3.3), (3.4) and $CAx = ACx$ we see that $\frac{d}{dt}\langle v(t), x^*\rangle = \langle v(t), A^*x^*\rangle$ for $x^* \in D(A^*)$ and $t \geq 0$. But our assumption in (iv) implies that this equation only has the zero solution. Hence $\int_0^t S(s)Ax = A\int_0^t S(s)xds$ for $t \geq 0$. This fact and (3.3) show that $x \in D(B)$ and $Bx = Ax$. Hence $A \subset B$.

"(i) $\Rightarrow$ (iv)" follows from Corollary 2.4 and the fact that $\|u(t; j_{\alpha-1}Cx, 0)\| = \|S(t)x\| \leq \|S(t)\|\|x\|$. Finally, we show "(iv) $\Rightarrow$ (i)". In view of Theorem 2.3, we need only to verify (iii) of Theorem 2.3. For any $x \in X$ let $\{x_m\}$ be a sequence in $D(A)$ such that $x_m \to x$. Let $u(\cdot; Cx_m)$ be the unique strong solution of the problem ACP$(j_{\alpha-1}Cx_m, 0)$, and let $v_m(t) = \int_0^t u(s; Cx_m)ds$. Then there is a continuous function u such that $u(t; Cx_m) \to u(t)$ and $v_m(t) \to v(t) = \int_0^t u(s)ds$ uniformly on compact subsets of $[0,\infty)$. Since A is closed and $Au(\cdot; Cx_m) = u'(\cdot; Cx_m) - j_{\alpha-1}(\cdot)Cx_m$ on $(0,\infty)$, we have

$$Av_m(t) = A\int_0^t u(s; Cx_m)ds = \int_0^t Au(s; Cx_m)ds = u(t; Cx_m) - j_\alpha(t)Cx_m,$$

which converges to $u(t) - j_\alpha(t)Cx$. It follows that $v(t) \in D(A)$ and $Av(t) = u(t) - j_\alpha(t)Cx = v'(t) - j_\alpha(t)Cx$. Hence v is a strong solution of $\mathrm{ACP}(j_\alpha Cx, 0)$. That this function v is the unique strong solution of $\mathrm{ACP}(j_\alpha Cx, 0)$ follows from the unique existence of the strong solution of $\mathrm{ACP}(0,0)$. Hence (iii) of Theorem 2.3 is satisfied. The proof is complete.

We end this section with the following characterization of α-times integrated semigroups, which is a specialization of Corollary 2.4 and Theorem 3.1.

Corollary 3.2. *Under the assumption that A is a densely defined closed operator with $\rho(A) \neq \emptyset$, the following statements are equivalent:*

(i) *A is the generator of an α-times integrated semigroup $S(\cdot)$ on X.*

(ii) *A_1 is the generator of an α-times integrated semigroup $S_1(\cdot)$ on X_1.*

(iii) $\mathrm{ACP}(j_{\alpha-1}x, 0)$ *has a unique strong solution for every $x \in D(A)$.*

(iv) $\mathrm{ACP}(j_\alpha x + j_\alpha * g, 0)$ *has a unique strong solution in $C^1([0,\infty), X)$ for every $g \in L^1_{loc}([0,\infty), X)$ and $x \in X$.*

(v) *For every $x \in D(A)$* $\mathrm{ACP}(j_{\alpha-1}x, 0)$ *has a unique strong solution which depends continuously on x.*

(vi) $\mathrm{ACP}(j_{\alpha-1} * g + j_{\alpha-1}x, 0)$ *has a unique weak solution for every $g \in L^1_{loc}([0,\infty), X)$ and $x \in X$.*

(vii) $\mathrm{ACP}(j_{\alpha-1}x, 0)$ *has a unique weak solution for every $x \in X$.*

4 Applications to Two Examples

For illustration, we consider two examples.

Example 1. Let $A := \sum_{j=0}^{k} a_j D^j$ be the maximal differential operator on a function space which can be any of the spaces

$$C_0(\mathbf{R}),\ C_b(\mathbf{R}),\ UC_b(\mathbf{R}),\ L^p(\mathbf{R}) \text{ for } 1 \le p \le \infty,$$

where $a_0, a_1, \ldots, a_k \in \mathbf{C}$ and $(D^j f)(x) = f^{(j)}(x), x \in \mathbf{R}$. It is shown in [12] that if the polynomial $p(x) := \sum_{j=0}^{k} a_j (ix)^j$ satisfies $w := \max\{0, \sup_{x\in\mathbf{R}} \mathrm{Re}(p(x))\} < \infty$, then, for $\alpha \in (\frac{1}{2}, 1]$, A generates an exponentially bounded, norm continuous α-times integrated semigroup $S(\cdot)$, which is defined by

$$(S(t)f)(x) := \frac{1}{\sqrt{2\pi}} (\tilde{\phi}_{\alpha,t} * f)(x)$$

with

$$\phi_{\alpha,t}(x) := \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} e^{p(x)s} ds,\ t \ge 0, x \in \mathbf{R}.$$

Here $\tilde{\phi}_{\alpha,t}$ denotes the inverse Fourier transform of $\phi_{\alpha,t}$.

An application of Corollary 2.4 shows that for every f in one of the spaces listed above, say $L^p(\mathbf{R})$ for example, and for every function g on $[0,\infty)\times\mathbf{R}$ which satisfies $\int_0^t(\int_{-\infty}^{\infty}|g(s,x)|^p dx)^{1/p}ds<\infty$ for all $t\in(0,\infty)$, the differential equation:

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) & = \sum\limits_{j=0}^{k} a_j(\frac{\partial}{\partial x})^j u(t,x)+\frac{t^\alpha}{\Gamma(\alpha+1)}f(x) \\ & \quad +\frac{1}{\Gamma(\alpha+1)}\int_0^t(t-s)^\alpha g(s,x)ds, t>0; \\ u(0,x) & = 0 \text{ a.e. } x\in\mathbf{R} \end{cases}$$

has a unique solution u, which is given by

$$u(t,x)=\frac{1}{\sqrt{2\pi}}\int_0^t\tilde{\phi}_{\alpha,s}*f(x)ds+\frac{1}{\sqrt{2\pi}}\int_0^t\int_0^s\tilde{\phi}_{\alpha,s-r}*g(r,x)drds$$

for $t\ge 0$, a.e. $x\in\mathbf{R}$.

Example 2. Let $A:=i(\Delta-V)$ be the Schrödinger operator on $L^p(\mathbf{R}^n), 1\le p<\infty, p\neq 2$, or $C_0(\mathbf{R}^n)$, or $BUC(\mathbf{R}^n)$, with potential $V\in L^\infty(\mathbf{R}^n)$. It is shown in [5, p. 77] that there exists $w\in\mathbf{R}$ such that A generates an $(w-\Delta+V)^{-r}$-group $\{e^{it(\Delta-V)}(w-\Delta+V)^{-r}\}_{t\in\mathbf{R}}$ on $L^p(\mathbf{R}^n)$ (resp. $C_0(\mathbf{R}^n)$ or $BUC(\mathbf{R}^n)$) for all $r>2n|\frac{1}{p}-\frac{1}{2}|$ (resp. $r>n$).

An application of Corollary 2.4 shows that for every f in one of the spaces listed above, say $L^p(\mathbf{R}^n)$ $(1\le p<\infty, p\neq 2)$ for example, and for every function g on $[0,\infty)\times\mathbf{R}^n$ which satisfies $\int_0^t(\int_{-\infty}^{\infty}|g(s,x)|^p dx)^{1/p}ds<\infty$ for all $t\in(0,\infty)$, the differential equation:

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) & = i(\Delta u(t,x)-V(x)u(t,x))+((w-\Delta+V)^{-r}f)(x) \\ & \quad +\int_0^t((w-\Delta+V)^{-r}g)(s,x)ds, t>0; \\ u(0,x) & = 0 \text{ a.e. } x\in\mathbf{R}^n \end{cases}$$

has a unique solution u, which is given by

$$\begin{aligned} u(t,x) &= \int_0^t\left(e^{is(\Delta-V)}(w-\Delta+V)^{-r}f\right)(x)ds \\ &\quad +\int_0^t\int_0^s\left(e^{i(s-\nu)(\Delta-V)}(w-\Delta+V)^{-r}g(\nu,\cdot)\right)(x)d\nu ds \end{aligned}$$

for $t\ge 0$, a.e. $x\in\mathbf{R}^n$.

Acknowledgement

The authors would like to thank the referees for their careful reading and valuable suggestions.

References

1. W. Arendt, Vector valued Laplace transforms and Cauchy problems, *Israel J. Math.* 59 (1987), 327-352.
2. J. M. Ball, Strongly continuous semigroups, weak solutions, and the variation of constants formula, *Proc. Amer. Math. Soc.* 63 (1977) 370-373.
3. E. B. Davies and M. M. Pang, The Cauchy problem and a generalization of the Hille-Yosida theorem, *Proc. London Math. Soc.* 55 (1987) 181-208.
4. R. deLaubenfels, C-semigroups and the Cauchy problem, *J. Funct. Anal.* 111 (1993) 44-61.
5. R. deLaubenfels, *Existence Families, Functional Calculi and Evolution Equations*, Lecture Notes in Mathematics Vol. 1570, Springer-Verlag Berlin Heidelberg 1994.
6. M. Hieber, Integrated semigroups and differential operators on L^p spaces, *Math. Ann.* 291 (1991), 1-16.
7. M. Hieber, Laplace transforms and α-times integrated semigroups, *Forum Math.* 3 (1991),595-612.
8. H. Kellerman and M. Hieber, Integrated semigroups, *J. Funct. Anal.* 84 (1989), 160-180.
9. C.-C. Kuo and S.-Y. Shaw, On α-times integrated C-semigroups and the abstract Cauchy problem, *Studia Math.* 142 (2000), 201-217.
10. Y.-C. Li and S.-Y. Shaw, N-times integrated C-semigroups and the abstract Cauchy problem, *Taiwanese J. Math.* 1 (1997), 75-102.
11. Y.-C. Li and S.-Y. Shaw, On generators of integrated C-semigroups and C-cosine functions, *Semigroup Forum* 47 (1993), 29-35.
12. M. Mijatović and S. Pilipović, α-times integrated semigroups ($\alpha \in \mathbf{R}^+$), *J. Math. Anal. Appl.* 210 (1997), 790-803.
13. I. Miyadera, M. Okubo and N. Tanaka, α-times integrated semigroups and abstract Cauchy problems, *Memoirs of the School of Science & Engineering, Waseda Univ.* 57 (1993), 267-289.
14. R. Nagel, *One Parameter Semigroups of Positive Operators*, Lecture Notes in Math., Vol. 1184, Springer-Verlag, New York/Berlin, 1986.
15. F. Neubrander, Integrated semigroups and their applications to the abstract Cauchy problem, *Pacific J. Math.* 135 (1988), 111-155.
16. N. Tanaka and I. Miyadera, Exponentially bounded C-semigroups and integrated semigroups, *Tokyo J. Math.* 12 (1989) 99-115.
17. N. Tanaka and I. Miyadera, C-semigroups and the abstract Cauchy problem, *J. Math. Anal. Appl.* 170 (1992) 196-206.
18. C.-C. Kuo and S.-Y. Shaw, Abstract Cauchy problems associated with local C-semigroups, in *Semigroups of Operators: Theory and Applications*, Proceedings of the Second International Conference, Rio de Janeiro, Brazil, September 10-14, 2001, Ed. by C. Kubrusly, N. Levan, and M. da Silveira, Optimization Software Inc., Publications, New York - Los Angeles, 2002, pp. 158-168.
19. S.-Y. Shaw and C.-C. Kuo, Generation of local C-semigroups and solvability of the abstract Cauchy problems, *Taiwanese J. Math.* 9 (2005), 291-311.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,213-232,
2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Best Approximation and Jackson-Type Estimates by Generalized Fuzzy Polynomials

Barnabás Bede and Sorin G. Gal
Department of Mathematics, University of Oradea,
Armatei Române 5, 410087 Oradea, Romania,
E-mail adresses: bede.barna@bgk.bmf.hu , galso@uoradea.ro

Abstract

In a very recent paper [3], it was proved that any 2π-periodic continuous fuzzy-number-valued function can be uniformly approximated by sequences of generalized fuzzy trigonometric polynomials, but without to give any estimate for the approximation error. In this paper, connected to a best approximation problem we obtain Jackson-type estimate. For the algebraic case we also obtain a Jackson-type estimate, using Szabados-type polynomials. Finally, as an application we study convergence of fuzzy Lagrange interpolation polynomials.

2000 AMS Subject Classification:26E50, 41A50, 41A05

Keywords and phrases: Fuzzy number, generalized fuzzy trigonometric polynomial, best apprximation, Jackson-type estimate, generalized fuzzy algebraic polynomial, fuzzy interpolation

1 Introduction

Firstly let us recall some known concepts and results. Let $\mathbb{R}_{\mathcal{F}}$ be the space of fuzzy real numbers (see e.g. [16]). For $0 < r \leq 1$ and $u \in\mathbb{R}_{\mathcal{F}}$ define $[u]^r = \{x \in \mathbb{R};\ u(x) \geq r\}$ and let $[u]^0 = \overline{\{x \in \mathbb{R};\ u(x) > 0\}}$. Then it is well-known that for each $r \in [0, 1]$, $[u]^r$ is a bounded closed interval, denoted by $[u]^r = [u^r_-, u^r_+]$, and for $u, v \in \mathbb{R}_{\mathcal{F}}$, $\lambda \in \mathbb{R}$, the sum $u \oplus v$ and the product $\lambda \odot u$ are defined by $[u \oplus v]^r = [u]^r + [v]^r$, $[\lambda \odot u]^r = \lambda[u]^r$, $\forall r \in [0, 1]$, where $[u]^r + [v]^r$ means the usual addition of two intervals (as subsets of $\mathbb{R}$) and $\lambda[u]^r$ means the usual product between a scalar and a subset of $\mathbb{R}$.

Defining $D:\mathbb{R}_{\mathcal{F}}\times\mathbb{R}_{\mathcal{F}}\rightarrow\mathbb{R}_{+}\cup\{0\}$ by $D(u,v)=\sup_{r\in[0,1]}\max\{|u_{-}^{r}-v_{-}^{r}|,|u_{+}^{r}-v_{+}^{r}|\}$, are well-known the following properties:
$D(u\oplus w,v\oplus w)=D(u,v)$, $\forall u,v,w\in\mathbb{R}_{\mathcal{F}}$,
$D(k\odot u,k\odot v)=|k|D(u,v)$, $\forall k\in\mathbb{R}$, $u,v\in\mathbb{R}_{\mathcal{F}}$,
$D(u\oplus v,w\oplus e)\leq D(u,w)+D(v,e)$, $\forall u,v,w,e\in\mathbb{R}_{\mathcal{F}}$
and $(\mathbb{R}_{\mathcal{F}},D)$ is a complete metric space.

A fuzzy-number-valued function $f:[a,b]\rightarrow\mathbb{R}_{\mathcal{F}}$ is called Riemann integrable to $I\in\mathbb{R}_{\mathcal{F}}$, if for any $\varepsilon>0$, there exists $\delta>0$ such that for any division $P=\{[u,v],\xi\}$ of $[a,b]$ with the norm $\Delta(P)<\delta$, we have

$$D\left(\sum_{\wp}(v-u)\odot f(\xi),I\right)<\varepsilon,$$

and we write $I=(R)\int_a^b f(x)dx$, (here $\sum$ is addition with respect to $\oplus$ in $\mathbb{R}_{\mathcal{F}}$).

A function $f:\mathbb{R}\rightarrow\mathbb{R}_{\mathcal{F}}$ will be caled $2\pi-$ periodic if $f(x+2\pi)=f(x)$, $\forall x\in\mathbb{R}$.

A generalized fuzzy trigonometric polynomial of degree $\leq n$ is defined as a finite sum of the form $T(x)=\sum_{k=0}^{n}t_k(x)\odot c_k$, where $c_k\in\mathbb{R}_{\mathcal{F}}$ and $t_k(x)$ are usual trigonometric polynomials of degree $\leq n$.

Let us denote $C_{2\pi}^{\mathcal{F}}(\mathbb{R})=\{f:\mathbb{R}\rightarrow\mathbb{R}_{\mathcal{F}};\ f$ is $2\pi-$periodic and continuous on $\mathbb{R}\}$.

In [3], the following Weierstrass-type result is proved.

Theorem 1.1. *For any $f\in C_{2\pi}^{\mathcal{F}}(\mathbb{R})$, there exists a sequence of generalized fuzzy trigonometric polynomials $(T_n(x))_{n\in\mathbb{N}}$ such that*

$$\lim_{n\rightarrow\infty}\sup_{x\in\mathbb{R}}D(T_n(x),f(x))=0.$$

Other results concerning approximation and interpolation of fuzzy-number-valued functions can be found in: [1], [7], [6], [10], [11], [14], [9]. But the problems of existence of best approximation fuzzy polynomials and of convergence of fuzzy Lagrange polynomials, were not yet considered by the fuzzy mathematical literature.

In Section 2 we consider some problems of best approximation by generalized fuzzy trigonometric polynomials (of degree $\leq n$) and a Jackson-type estimate is proved.

Section 3 contains the case of best approximation by generalized fuzzy algebraic polynomials.

In Section 4, as an application, we prove the convergence of Lagrange interpolating polynomials for the class of fuzzy Lipschitz functions of order $> \frac{1}{2}$.

2 Best approximation, trigonometric case

On $C_{2\pi}^{\mathcal{F}}(\mathbb{R})$ let us consider the uniform distance

$$D^*(f,g) = \sup\{D(f(x), g(x); x \in \mathbb{R}\} = \sup\{D(f(x), g(x); x \in [-\pi, \pi]\},$$

$\forall f, g \in C_{2\pi}^{\mathcal{F}}(\mathbb{R})$.

For an interval $I \subset \mathbb{R}$ and a subset $K \subset \mathbb{R}_{\mathcal{F}}$, $K \neq \emptyset$, let us consider

$V_n^{K,I} = \{T_n; T_n(x) = \sum_{k=0}^n t_k(x) \odot c_k$, where all $c_k \in K$ and each $t_k(x)$ is an usual trigonometric polynomial of degree $\leq n$ with all its coefficients belonging to $I\}$.

For fixed $f \in C_{2\pi}^{\mathcal{F}}(\mathbb{R})$, $K \subset \mathbb{R}_{\mathcal{F}}$ and $I \subset \mathbb{R}$ and for each $n \in \mathbb{N}$, it is natural to consider the following problem of best approximation.

$E_n^{K,I}(f) = \inf\{D^*(f, T_n); T_n \in V_n^{K,I}\}$.

In the study of this problem, it is essential the following

Theorem 2.1. *If $K \subset \mathbb{R}_{\mathcal{F}}$, $K \neq \emptyset$, is a compact and $I = [A, B]$ is compact subinterval of $\mathbb{R}$, then the set $V_n^{K,I}$ is sequentially compact in the metric space $(C_{2\pi}^{\mathcal{F}}(\mathbb{R}), D^*)$, for all $n \in \mathbb{N}$.*

Proof. Let us denote by $\mathcal{T}_n^I = \{t_k; t_k$ is usual trigonometric polynomial of degree $\leq n$, with all coefficients belonging to $I\}$ and define $\varphi : K^{n+1} \times (\mathcal{T}_n^I)^{n+1} \to C_{2\pi}^{\mathcal{F}}(\mathbb{R})$, by $\varphi(c_0, ..., c_n, t_0, ..., t_n)(x) = \sum_{k=0}^n t_k(x) \odot c_k$.

Firstly let us prove that φ is continuous.

Indeed, let another generalized fuzzy trigonometric polynomial of degree $\leq n$, $\sum_{k=0}^n s_k(x) \odot d_k$. By the properties of D in Introduction and by [3, Lemma 2.2], we get

$$D\left(\sum_{k=0}^n t_k(x) \odot c_k, \sum_{k=0}^n s_k(x) \odot d_k\right) \leq$$

$$\sum_{k=0}^n |t_k(x)| D(c_k, d_k) + \sum_{k=0}^n |t_k(x) - s_k(x)| \odot D(c_k, \widetilde{0}),$$

where $\widetilde{0} = \chi_{\{0\}} \in \mathbb{R}_{\mathcal{F}}$.

Because each $t_k(x)$ is of the form $\alpha_0 + \sum_{j=0}^{n}(\alpha_j \cos jx + \beta_j \sin jx)$, with $\alpha_j, \beta_j \in I = [A, B]$, it immediately follows that

$|t_k(x)| \leq (2n+1)\max\{|A|, |B|\} = M$, for all $k = \overline{0,n}$ and all $x \in \mathbb{R}$.

Also, because $c_k \in K-$ compact, $\forall k = \overline{0,n}$, we get that $K' = K \cup \{\widetilde{0}\}$ is compact too (in the metric space $\mathbb{R}_{\mathcal{F}}$), which implies that it is bounded and therefore there exists a constant $M' > 0$ such that $D(c_k, \widetilde{0}) \leq M$, $\forall c_k \in K$.

As a conclusion, it follows

$$D\left(\sum_{k=0}^{n} t_k(x) \odot c_k, \sum_{k=0}^{n} s_k(x) \odot d_k\right) \leq$$

$$M \sum_{k=0}^{n} D(c_k, d_k) + M' \sum_{k=0}^{n} \|t_k(x) - s_k(x)\|$$

(here $\|\cdot\|$ denotes the usual uniform norm on the set of real valued, $2\pi-$ periodic functions, denoted by $C_{2\pi}$).

This last inequality immediately shows that φ is continuous, if $K^{n+1} \times (\mathcal{T}_n^I)^{n+1}$ is endowed with the box metric given by

$$\rho[(c_0, ..., c_n, t_0, ..., t_n), (d_0, ..., d_n, s_0, ..., s_n)] = \max_{k=\overline{0,n}}\{D(c_k, d_k), \|t_k - s_k\|\}.$$

Now we claim that $\mathcal{T}_n^I$ is compact in $(C_{2\pi}, \|\cdot\|)$. Indeed, if we consider $\psi : I^{2n+1} \to C_{2\pi}$ defined by

$$\psi(\alpha_0, ..., \alpha_n, \beta_1, ..., \beta_n)(x) = \alpha_0 + \sum_{k=0}^{n}[\alpha_k \cos kx + \beta_k \sin kx],$$

then it easily follows that ψ is continuous and therefore $\mathcal{T}_n^I = \psi(I^{2n+1})$ is compact.

As a conclusion, $K^{n+1} \times (\mathcal{T}_n^I)^{n+1}$ is compact which implies that $V_n^{K,I} = \varphi(K^{n+1} \times (\mathcal{T}_n^I)^{n+1})$ is compact, and therefore as a compact subset of a metric space, $V_n^{K,I}$ is sequentially compact. □

As an immediate consequence of Theorem 2.1, we get

Corollary 2.2. *Let $f \in C_{2\pi}^{\mathcal{F}}(\mathbb{R})$. If $K \subset \mathbb{R}_{\mathcal{F}}$, $K \neq \emptyset$ is compact and $I = [A, B]$ is a compact interval of $\mathbb{R}$, then for each $n \in \mathbb{N}$, there exists $T^* \in V_n^{K,I}$ such that $E_n^{K,I}(f) = D^*(f, T^*)$, i.e. T^* is a generalized fuzzy trigonometric polynomial (of degree $\leq n$) of best approximation for f.*

Proof. Since $K \neq \emptyset$, it follows that $V_n^{K,I} \neq \emptyset$.

For $\varepsilon = \frac{1}{m}$, there exists $T_m \in V_n^{K,I}$ such that $E_n^{K,I}(f) \leq D^*(f, T_m) \leq E_n^{K,I}(f) + \frac{1}{m}$, $m = 1, 2, \ldots$. Since by Theorem 2.1, $V_n^{K,I}$ is sequentially compact, the sequence $(T_m)_m$ has a convergent subsequence $(T_{m_k})_k$ to an element $T^* \in V_n^{K,I}$. Passing above to limit, we get $E_n^{K,I}(f) = D^*(f, T^*)$, i.e. T^* is of best approximation, which proves the corollary. □

Remark 2.3. If $f \in C_{2\pi}^{\mathcal{F}}(\mathbb{R})$, then by $f([-\pi, \pi]) = K$ compact, it follows that in Corollary 2.2 we can take $K = f([-\pi, \pi])$ (i.e. depending on f).

In what follows we will derive a Jackson-type estimate for $E_n^{K,I}(f)$ with $K = f([-\pi, \pi])$.

Theorem 2.4. *If $f \in C_{2\pi}^{\mathcal{F}}(\mathbb{R})$ and $[-1, 1] \subset [A, B]$, then there exists a constant $C > 0$ (independent of f and n) and an index $n_0 \in \mathbb{N}$ (independent of f) such that for $K = f([-\pi, \pi])$ we have*

$$E_n^{K,[A,B]}(f) \leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right), \ \forall n \geq n_0$$

where $\omega_1^{\mathcal{F}}(f, \delta) = \sup\{D(f(x), f(y); x, y \in \mathbb{R}, |x - y| \leq \delta\}$.

Proof. In [8, p. 646] was introduced the following fuzzy Jackson operator

$$J_n(f)(x) = (R)\int_{-\pi}^{\pi} K_n(t) \odot f(x + t)dt = (R)\int_{-\pi}^{\pi} K_n(u - x) \odot f(u)du,$$

where $K_n(t) = L_{n'}(t)$, $n' = [n/2] + 1$,

$$L_{n'}(t) = \frac{3}{2\pi n'[2(n')^2 + 1]}\left[\frac{\sin(n't/2)}{\sin(t/2)}\right]^4, \quad \int_{-\pi}^{\pi} L_{n'}(t)dt = 1,$$

and it was proved [8, p.647, Theor. 13.14] the estimate

$$D(J_n(f)(x), f(x) \leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right), \ \forall n \in \mathbb{N}, \ x \in \mathbb{R}.$$

On the other hand by taking the Riemann sum of $J_n(f)(x)$ (on an equidistant division of $[-\pi, \pi]$), we get

$$T_n(x) = \frac{2\pi}{n'}\sum_{k=0}^{n'} L_{n'}\left(-\pi + \frac{2k\pi}{n'} - x\right) \odot f\left(-\pi + \frac{2k\pi}{n'}\right).$$

Obviously $T_n(x)$ is a generalized fuzzy trigonometric polynomial of degree $\leq n$ (since $n' \leq n$) and by [4, Corollary 3], for all $x \in [-\pi, \pi]$ and $n \in \mathbb{N}$, we have

$$D(J_n(f)(x), T_n(x)) \leq 2\pi\omega_1^{\mathcal{F}}\left(f, \frac{2\pi}{n'}\right)_{[-\pi,\pi]}$$

$$\leq 2\pi(2\pi+1)\omega_1^{\mathcal{F}}\left(f, \frac{1}{n'}\right)_{[-\pi,\pi]} \leq 4\pi(2\pi+1)\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{[-\pi,\pi]}$$

(since $n' = [n/2] + 1 > n/2$).

But reasoning exactly as in the usual case (see [2, p.75, Lemma 2.2.1], we have $\omega_1^{\mathcal{F}}(f,\delta)_{[-\pi,\pi]} \leq \omega_1^{\mathcal{F}}(f,\delta) \leq 2\omega_1^{\mathcal{F}}(f,\delta)_{[-\pi,\pi]}$.

As a consequence, we obtain the following Jackson-type estimate

$$D(T_n(x), f(x)) \leq D(T_n(x), J_n(f)(x)) + D(J_n(f)(x), f(x)) \leq$$

$$\leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right), \text{ for all } n \in \mathbb{N},\ x \in [-\pi, \pi].$$

(Note that above $\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)$ can be replaced by $\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{[-\pi,\pi]}$ too).

To finish the proof, we have to calculate the bounds for the coefficients of the usual trigonometric polynomials $L_{n'}\left(-\pi + \frac{2k\pi}{n'} - x\right)$ in the expression of $T_n(x)$.

Firstly, it is well known the identity

$$\left(\frac{\sin\frac{mx}{2}}{\sin\frac{x}{2}}\right)^2 = m + 2\sum_{k=1}^{m-1}(m-k)\cos kx.$$

It follows

$\left(\frac{\sin\frac{mx}{2}}{\sin\frac{x}{2}}\right)^4 = \left(m + 2\sum_{k=1}^{m-1}(m-k)\cos kx\right)\left(m + 2\sum_{k=1}^{m-1}(m-k)\cos kx\right) =$
$m^2 + 4m\sum_{k=1}^{m-1}(m-k)\cos kx + 4\sum_{i=1}^{m-1}\sum_{j=1}^{m-1}(m-i)(m-j)\cos(ix)\cos(jx) =$
$= m^2 + 4m\sum_{k=1}^{m-1}(m-k)\cos kx + 4\sum_{i=1}^{m-1}\sum_{j=1}^{m-1}(m-i)(m-j)\{\frac{1}{2}[\cos x(i+j) + \cos x(i-j)]\} =$
$= m^2 + 4m\sum_{k=1}^{m-1}(m-k)\cos kx + 2\sum_{i=1}^{m-1}\sum_{j=1}^{m-1}(m-i)(m-j)\cos x(i+j) + 2\sum_{i=1}^{m-1}\sum_{j=1}^{m-1}(m-i)(m-j)\cos x(i-j) =$
$= m^2 + 4m\sum_{k=1}^{m-1}(m-k)\cos kx + 2\sum_{\substack{i,j=1\\ i+j\leq m}}^{m-1}(m-i)(m-j)\cos x(i+j) + 2\sum_{\substack{i,j=1\\ i+j>m}}^{m-1}(m-i)(m-j)\cos x(i+j) + 2\sum_{i,j=1}^{m-1}(m-i)(m-j)\cos x(i-j) := m^2 + S_1 + S_2 + S_3 + S_4.$

By simple calculation we can write

$$S_2 = \sum_{k=2}^{m} [2 \sum_{i=1}^{k-1} (m-i)(m-(k-i))] cos(kx),$$

$$S_3 = \sum_{p=1}^{m-2} [2 \sum_{i=1}^{m-p-1} (m-(p+i))(m-(m-i))] cos((m+p)x),$$

$$S_4 = 4(1^2 + 2^2 + ... + (m-1)^2) + \sum_{k=1}^{m-2} [4 \sum_{i=1}^{m-1-k} (m-i)(m-(k+i))] cos(kx).$$

By the relations

$$\sum_{i=1}^{k-1} (m-i)(m-(k-i)) = \sum_{i=1}^{k-1} [-i^2 + ik + m(m-k)] \leq \sum_{i=1}^{k-1} [k^2 + 4m(m-k)]/2 \leq$$

$$(k-1)[k^2 + 4m(m-k)]/2 \leq (m-1)[m^2 + 4m(m-2)]/2, k = 2, ..., m,$$

$$\sum_{i=1}^{m-p-1} (m-(p+i))(m-(m-i)) = \sum_{i=1}^{m-p-1} [-i^2 + i(m-p)] \leq \sum_{i=1}^{m-p-1} (m-p)^2/2 \leq$$

$$(m-p)^2(m-p-1)/2 \leq (m-1)^2(m-2)/2, p = 1, ..., m-2,$$

$$\sum_{i=1}^{m-1-k} (m-i)(m-i-k) \leq \sum_{i=1}^{m-1-k} (m-1)(m-1-k) \leq$$

$$(m-1)(m-1-k)^2 \leq (m-1)(m-2)^2, k = 1, ..., m-2,$$

$$1^2 + 2^2 + ... + (m-1)^2 = m(m-1)(2m-1)/6,$$

it follows that for $k \in \{0, ..., 2m-2\}$, the coefficients of $\cos kx$ in $\left(\frac{\sin \frac{mx}{2}}{\sin \frac{x}{2}}\right)^4$ are all positive and bounded by an algebraic polynomials of degree 3, with constant coefficients, independent of f, let us denote it by $H_3(m)$ (in S_1, obviously all the coefficients of $cos(kx)$ are bounded by $4m(m-1)$).

As a conclusion, it easily follows that in $(2\pi/n')L_{n'}\left(-\pi + \frac{2k\pi}{n'} - x\right)$ (which contains terms in $\cos kx$ and $\sin kx$), all the coefficients are bounded, in absolute value, by $F(n') = 3H_3(n')/[(n')^2(2(n')^2+1)]$, that is an $n_0 \in \mathbb{N}$ (independent of f) can be found (constructively), such that for all $n' \geq n_0$ we have $F(n') \leq 1$ (since $F(n')$ converges to 0 when n' converges to infinity).

Therefore, for $n \geq 3n_0$, it follows that $T_n(x)$ belongs to $V_n^{K,[-1,1]}$. Now, for $[-1,1] \subset [A,B]$ it is obvious that $E_n^{K,[A,B]}(f) \leq E_n^{K,[-1,1]}(f)$, which proves the theorem. □

Remark 2.5. ¿From the proof it is easily seen that an interval $[A,B]$ (independent of f and n) can be constructively determined such that the Jackson kind estimate in Theorem 2.4 holds for all $n = 1, 2, ...$.

3 Best approximation, algebraic case

Let $C_{\mathcal{F}}[a,b] = \{f : [a,b] \to \mathbb{R}_{\mathcal{F}}; f$ continuous on $[a,b]\}$ where $[a,b]$ is a compact subinterval of $\mathbb{R}$. If we define the concept of generalized fuzzy algebraic polynomial of degree $\leq n$ as in [17], i.e. as a finite sum of the form $\sum_{k=0}^{n} p_k(x) \odot c_k$, where $c_k \in \mathbb{R}_{\mathcal{F}}$ and $p_k(x)$ are algebraic polynomials of degree $\leq n$, we can repeat the reasonings in the above Theorem 2.1 and Corollary 2.2 simply by replacing $[-\pi, \pi]$ there by $[a,b]$, $C_{2\pi}^{\mathcal{F}}(\mathbb{R})$ by $C_{\mathcal{F}}[a,b]$ and the generalized fuzzy trigonometric polynomials by generalized fuzzy algebraic polynomials. But if we consider probably the simplest generalized fuzzy algebraic polynomials, given by the fuzzy Bernstein polynomials

$$B_n(f)(x) = \sum_{k=0}^{n} p_{n,k}(x) \odot f(k/n), x \in [0,1]$$

where $p_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$, we easily see that the coefficients of x^s in $p_{n,k}(x)$ are in general unbounded, for $k = \overline{1, n-1}$, while however $|p_{n,k}(x)| \leq 1$, $\forall x \in [0,1]$, $n \in \mathbb{N}$, $k = \overline{0,n}$. Therefore, in algebraic case, would be more natural to consider the problem of best approximation as follows. For a constant $M > 0$ and a subset $K \subset \mathbb{R}_{\mathcal{F}}$, $K \neq \emptyset$, let us consider $A_n^{K,M}[a,b] = \{P_n;\ P_n(x) = \sum_{k=0}^{n} p_k(x) \odot c_k$, where all $c_k \in \mathbb{R}_{\mathcal{F}}$ and $p_k(x)$ are algebraic polynomials of degree $\leq n$, satisfying $|p_k(x)| \leq M$, for all k and all $x \in [a,b]\}$.

For fixed $f \in C_{\mathcal{F}}[a,b]$, $K \subset \mathbb{R}_{\mathcal{F}}$, $K \neq \emptyset$ and $M > 0$ and for each $n \in \mathbb{N}$, we can consider the following problem of best approximation

$$E_n^{K,M}(f) = \inf\{D^*(f, P_n); P_n \in A_n^{K,M}[a,b]\},$$

where $D^*(f,g) = \sup\{D(f(x), g(x);\ x \in [a,b]\}$, for $f, g \in C_{\mathcal{F}}[a,b]$.

We have:

Theorem 3.1. *If $K \subset \mathbb{R}_{\mathcal{F}}$, $K \neq \emptyset$ is compact and $M > 0$, then the set $A_n^{K,M}[a,b]$ is sequentially compact in the metric space $(C_{\mathcal{F}}[a,b], D^*)$, for all $n \in \mathbb{N}$.*

Proof. Let us denote by $\mathcal{P}_n^M = \{p; p$ usual algebraic polynomials of degree $\leq n$, satisfying $|p(x)| \leq M$, for all $x \in [a,b]\}$ and define $\varphi : K^{n+1} \times (\mathcal{P}_n^M)^{n+1} \to C_{\mathcal{F}}[a,b]$, by $\varphi(c_0, ..., c_n, p_0, ..., p_n)(x) = \sum_{k=0}^{n} p_k(x) \odot c_k$.

Reasoning exactly as in the proof of Theorem 2.1, we get that φ is continuous and because $\mathcal{P}_n^M$ is compact in $C[a,b]$ (endowed with the uniform norm $\|\cdot\|$), see e.g. [13, p. 16 Lemma 1], we get the desired conclusion. □

Consequently, we obtain the following

Corollary 3.2. *Let $f \in C_{\mathcal{F}}[a,b]$. If $K \subset \mathbb{R}_{\mathcal{F}}$, $K \neq \emptyset$ is compact and $M > 0$, then for all $n \in \mathbb{N}$, there exists $T^* \in A_n^{K,M}[a,b]$ such that $E_n^{K,M}(f) = D^*(f, T^*)$, i.e. T^* is a generalized fuzzy algebraic polynomial (of degree $\leq n$), of best approximation for f.*

Remark 3.3. By [8, p.642, Theorem 13.13], we immediately abtain

$$E_n^{K,1}(f) \leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{\sqrt{n}}\right)_{[0,1]}, \ \forall n \in \mathbb{N},\ f \in C_{\mathcal{F}}[0,1],\ K = f([0,1]).$$

In what follows we deduce Jackson-type estimate for $E_n^{K,M}(f)$ by using some fuzzy analogous of Szabados-type polynomials (see e.g.[15]). For this aim we need the following lemmas.

Lemma 3.4. *Let $f : \left[-\frac{1}{4}, \frac{1}{4}\right] \to \mathbb{R}_{\mathcal{F}}$, be continuous and*

$$R_n(f,x) = \sum_{k=-n}^{n} r_{n,k}(x) \odot f(x_k),$$

with $r_{n,k}(x) = \frac{(x-x_k)^{-4}}{\sum_{j=-n}^{n}(x-x_j)^{-4}}$ and $x_k = \frac{k}{4n}$, $k = \overline{-n,n}$ Then the following estimate holds true:

$$D^*(f, R_n) \leq 5\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{\left[-\frac{1}{4},\frac{1}{4}\right]}.$$

Proof. We follow the proof of Lemma 1 in [15] for $r = 0$ and $s = 4$. Thus, for fixed x, let i be an index such that

$$|x - x_i| = \min_{|k| \leq n} |x - x_k| \leq \frac{1}{8n}. \tag{1}$$

Then evidently

$$\frac{|i-k|}{8n} \leq |x - x_k| \leq \frac{|i-k|}{2n}, \text{ for } i \neq k. \tag{2}$$

Denote $I = \left[-\frac{1}{4}, \frac{1}{4}\right]$. Since $r_{n,k}(x) \geq 0$, $\forall k = \overline{-n, n}$ and $\sum_{k=-n}^{n} r_{n,k}(x) = 1$, by the properties of D we have:

$$D(f(x), R_n(f,x)) = D\left(\left(\sum_{k=-n}^{n} r_{n,k}(x)\right) \odot f(x), \sum_{k=-n}^{n} r_{n,k}(x) \odot f(x_k)\right)$$

$$\leq \sum_{k=-n}^{n} r_{n,k}(x) \odot D(f(x), f(x_k)) \leq \sum_{k=-n}^{n} r_{n,k}(x) \omega_1^{\mathcal{F}} (f, |x - x_k|)_I$$

$$\leq (x - x_i)^4 \sum_{k=-n}^{n} \frac{D(f(x), f(x_k))}{(x - x_k)^4} \leq (x - x_i)^4 \sum_{k=-n}^{n} |x - x_k|^{-4} \omega_1^{\mathcal{F}} (f, |x - x_k|)_I$$

$$\leq \omega_1^{\mathcal{F}} (f, |x - x_i|)_I + (8n)^{-4} \sum_{\substack{k=-n \\ k \neq i}}^{n} \omega_1^{\mathcal{F}} \left(f, \frac{|i-k|}{2n}\right)_I \left(\frac{8n}{|i-k|}\right)^4$$

$$\leq \left[1 + \sum_{\substack{k=-n \\ k \neq i}}^{n} |i-k|^{-2}\right] \omega_1^{\mathcal{F}} \left(f, \frac{1}{n}\right)_I \leq 5\omega_1^{\mathcal{F}} \left(f, \frac{1}{n}\right)_I.$$

□

Remark 3.5. The result in the above Lemma 3.4 can be seen as a Jackson-type estimate for the error of the approximation by fuzzy generalized rational functions.

For the next results we need an embedding theorem.

Theorem 3.6. *(see e.g. [16]) Let* $\overline{C}[0,1]$ *be the class of all real valued bounded functions* f *on* $[0,1]$*, such that* f *is left continuous on* $(0,1]$ *and* f *has right limit for* $x \in [0,1)$*, especially* f *is right continuous at* 0*. With the norm* $\|f\| = \sup_{x\in[0,1]} |f(x)|$, $\overline{C}[0,1]$ *is a Banach space. For* $u \in \mathbb{R}_{\mathcal{F}}$*, define* $j : \mathbb{R}_{\mathcal{F}} \to \overline{C}[0,1]$, $j(u) = (u_-, u_+)$*, where* $u_-(r) = u_-^r$ *and* $u_+(r) = u_+^r$*. Then* $j(\mathbb{R}_{\mathcal{F}})$ *is a closed convex cone in the Banach space* $\overline{C}[0,1] \times \overline{C}[0,1]$ *and:*

(i) $j(s \odot u \oplus t \odot v) = s \cdot j(u) + t \cdot j(v)$, $\forall u, v \in \mathbb{R}_{\mathcal{F}}$ *and* $s, t \in \mathbb{R}_+$ *(here "+" and "·"denote the addition and scalar multiplication in* $\overline{C}[0,1] \times \overline{C}[0,1]$*).*

(ii) $D(u,v) = \|j(u) - j(v)\|$, $\forall u, v \in \mathbb{R}_{\mathcal{F}}$*. i.e.* j *embeds* $\mathbb{R}_{\mathcal{F}}$ *in* $\overline{C}[0,1] \times \overline{C}[0,1]$ *isometrically and isomorphically (*$\|\cdot\|$ *beeing the usual product norm in* $\overline{C}[0,1] \times \overline{C}[0,1]$*).*

The following Lemmas give some approximation properties in Banach spaces.

Lemma 3.7. *Let* $(\mathbb{B}, \|\cdot\|)$ *be a Banach space and* $g : \left[-\frac{1}{4}, \frac{1}{4}\right] \to \mathbb{B}$ *continuous. Let* $R_n(g,x) = \sum_{k=-n}^{n} r_{n,k}(x) \cdot g(x_k)$*, with* $r_{n,k}$ *as in Lemma 3.4. Then*

$$\|R_n'(g,x)\| \leq 900 n \omega_1^{\mathbb{B}} \left(g, \frac{1}{n}\right)_{\left[-\frac{1}{4}, \frac{1}{4}\right]},$$

where $R_n'(g,x)$ *is the Fréchet derivative of* $R_n(g,x)$ *in* $\mathbb{B}$ *and* $\omega_1^{\mathbb{B}}(g,\delta)_{\left[-\frac{1}{4},\frac{1}{4}\right]} = \sup\left\{\|g(x) - g(y)\|;\ x, y \in \left[-\frac{1}{4}, \frac{1}{4}\right],\ |x - y| \leq \delta\right\}.$

Proof. The proof is the same as that of [15, Lemma 2], written in the case of functions with values in a Banach space.

Thus by (1) and (2) we get

$$\|R_n'(g,x)\| = \left\| \frac{-4\sum_{k=-n}^{n} g(x_k)(x-x_k)^{-5} \sum_{k=-n}^{n} (x-x_k)^{-4}}{\left(\sum_{k=-n}^{n}(x-x_k)^{-4}\right)^2} + \right.$$

$$\left. + \frac{4\sum_{k=-n}^{n} g(x_k)(x-x_k)^{-4} \sum_{k=-n}^{n} (x-x_k)^{-5}}{\left(\sum_{k=-n}^{n}(x-x_k)^{-4}\right)^2} \right\|$$

$$= 2\frac{\left\|\sum_{k=-n}^{n}(x-x_k)^{-5}\sum_{j=-n}^{n}[g(x_k)-g(x_j)](x_k-x_j)(x-x_j)^{-5}\right\|}{\left(\sum_{k=-n}^{n}(x-x_k)^{-4}\right)^2}$$

$$\leq 2(x-x_i)^8 \sum_{k=-n}^{n} |x-x_k|^{-5} \cdot \frac{1}{4n}\omega_1^{\mathbb{B}}\left(g, \frac{1}{n}\right)_{\left[-\frac{1}{4},\frac{1}{4}\right]} \sum_{\substack{j=-n \\ j\neq k}}^{n} \frac{|j-k|^2}{|x-x_j|^5}$$

$$\leq \frac{(8n)^{-3}\omega_1^{\mathbb{B}}\left(g,\frac{1}{n}\right)_{\left[-\frac{1}{4},\frac{1}{4}\right]}}{2n}\cdot$$

$$\cdot\left\{\sum_{\substack{j=-n\\ j\neq i}}^{n}\left(\frac{8n}{|j-i|}\right)^5|j-i|^2+\sum_{\substack{k=-n\\ k\neq i}}^{n}\left(\frac{8n}{|k-i|}\right)^5\left[|k-i|^2+\sum_{\substack{j=-n\\ j\neq i}}^{n}\frac{|j-k|^2}{|j-i|^5}\right]\right\}$$

$$\leq 32n\omega_1^{\mathbb{B}}\left(g,\frac{1}{n}\right)_{\left[-\frac{1}{4},\frac{1}{4}\right]}\left\{\sum_{\substack{j=-n\\ j\neq i}}^{n}|j-i|^{-3}+\sum_{\substack{k=-n\\ k\neq i}}^{n}|k-i|^{-3}\left[1+4\sum_{\substack{j=-n\\ j\neq i}}^{n}|j-i|^{-3}\right]\right\}$$

$$\leq 900n\omega_1^{\mathbb{B}}\left(g,\frac{1}{n}\right)_{\left[-\frac{1}{4},\frac{1}{4}\right]}.$$

□

Lemma 3.8. *Let* $(\mathbb{B},\|\cdot\|)$ *be a Banach space and* $f:\left[-\frac{1}{2},\frac{1}{2}\right]\to\mathbb{B}$ *continuous. Let* $P_n(x)=c_n\left(\frac{\cos(2n\arccos x)}{x^2-\sin^2\frac{\pi}{4n}}\right)^2$ *where* c_n *is choosen such that* $\int_{-1}^{1}P_n(x)dx=1$ *and let*

$$K_n(f,x)=\int_{-\frac{1}{2}}^{\frac{1}{2}}[f(t)-f(0)]P_n(t-x)dt+f(0)$$

be the Bojanic-DeVore operator, where the integral is considered to be the usual Riemann integral for functions $g:\left[-\frac{1}{2},\frac{1}{2}\right]\to\mathbb{B}$*. Then*

$$\|f-K_n(f)\|_{C\left(\left[-\frac{1}{4},\frac{1}{4}\right],\mathbb{B}\right)}\leq C_4\omega_1^{\mathbb{B}}\left(f,\frac{1}{n}\right)_{\left[-\frac{1}{4},\frac{1}{4}\right]}.$$

Proof. The proof is the same as the proof of [13, p. 275-276, Proposition 3.4] but for functions with values in a Banach space. Indeed, firstly $P_n(x)$ is an even algebraic polynomial of degree $4n-4$, therefore $K_n(f,x)$ is a generalized (algebraic) polynomial of degree $4n-4$, with coefficients in the Banach space $\mathbb{B}$. Let us denote $I=[-1,1]$, $I'=\left[-\frac{1}{4},\frac{1}{4}\right]$ and for fixed $x\in I'$, $I_x=\left[-\frac{1}{2}+x,\frac{1}{2}-x\right]$. If we denote $g(x)=f(x)-f(0)$ and $L_n(g,x)=\int_{-\frac{1}{2}}^{\frac{1}{2}}g(t)P_n(t-x)dt$, then by [13, p.276, relation (3.10)], it follows

$$\|f-K_n(f)\|_{C(I',\mathbb{B})}=\|g-K_n(g)\|_{C(I',\mathbb{B})}, \tag{3}$$

where $C(I', \mathbb{B}) = \{f : I' \to \mathbb{B};\ f \text{ continuous on } I'\}$. Then for fixed $x \in I'$, as in [13, p.276] we get

$$L_n(g,x) - g(x) = \int_{I_x} [g(x+u) - g(x)] P_n(u) du - g(x) \int_{I \setminus I_x} P_n(u) du$$

where $\left\| g(x) \int_{I\setminus I_x} P_n(u) du \right\|_{C(I',\mathbb{B})} \leq C_2 \|g\|_{C(I',\mathbb{B})} \cdot n^{-2}$ and

$$\left\| \int_{I_x} [g(x+u) - g(x)] P_n(u) du \right\|_{C(I',\mathbb{B})} \leq C_1 \omega_1^{\mathbb{B}} \left(g, \frac{1}{n}\right)_{I'} = C_1 \omega_1^{\mathbb{B}} \left(f, \frac{1}{n}\right)_{I'}.$$

Since $\|g\|_{C(I',\mathbb{B})} = \|f - f(0)\|_{C(I',\mathbb{B})} \leq \omega_1^{\mathbb{B}} \left(f, \frac{1}{4}\right)_{I'}$, by (3) we obtain

$$\|f - K_n(f)\|_{C(I',\mathbb{B})} \leq C_3 \left[\omega_1^{\mathbb{B}} \left(f, \frac{1}{n}\right)_{I'} + \omega_1^{\mathbb{B}} \left(f, \frac{1}{4}\right)_{I'} \cdot n^{-2}\right] \leq C_4 \omega_1^{\mathbb{B}} \left(f, \frac{1}{n}\right)_{I'},$$

taking into account that $\omega_1^{\mathbb{B}} \left(f, \frac{1}{4}\right)_{I'} \leq \omega_1^{\mathbb{B}} (f, 1)_{I'} = \omega_1^{\mathbb{B}} \left(f, n \cdot \frac{1}{n}\right)_{I'} \leq$
$\leq n\omega_1^{\mathbb{B}} \left(f, \frac{1}{n}\right)_{I'} \leq n^2 \omega_1^{\mathbb{B}} \left(f, \frac{1}{n}\right)_{I'}$. The lemmma is proved. □

Now let us consider $h : \left[-\frac{1}{4}, \frac{1}{4}\right] \to \mathbb{B}$ and

$$\overline{R}_n(h,x) = \begin{cases} h\left(-\frac{1}{4}\right) \text{ if } -\frac{1}{2} \leq x \leq -\frac{1}{4} \\ R_n(h,x) \text{ if } -\frac{1}{4} \leq x \leq \frac{1}{4} \\ h\left(\frac{1}{4}\right) \text{ if } \frac{1}{4} \leq x \leq \frac{1}{2} \end{cases}.$$

For $f \in C_{\mathcal{F}}\left[-\frac{1}{4}, \frac{1}{4}\right]$, let $K_n(\overline{R}_n(j \circ f), x)$ be the Bojanic-DeVore operator associated to $\overline{R}_n(j \circ f)$, where j is the embedding in Theorem 3.6, i.e.

$$K_n(\overline{R}_n(j \circ f), x) = \int_{-\frac{1}{2}}^{\frac{1}{2}} [\overline{R}_n(j \circ f)(t) - \overline{R}_n(j \circ f)(0)] P_n(t-x) dt +$$

$$\overline{R}_n(j \circ f)(0).$$

Then we have:

$$K_n(\overline{R}_n(j \circ f), x) = (j \circ f)(0) \left[1 - \int_{-\frac{1}{2}}^{\frac{1}{2}} P_n(t-x) dt\right]$$

$$+ (j \circ f)\left(-\frac{1}{4}\right) \int_{-\frac{1}{2}}^{-\frac{1}{4}} P_n(t-x) dt$$

$$+\sum_{k=-n}^{n}(j\circ f)(x_k)\int_{-\frac{1}{4}}^{\frac{1}{4}}\frac{P_n(t-x)dt}{(t-x_k)^4\sum_{j=-n}^{n}(t-x_j)^{-4}}$$
$$+(j\circ f)\left(\frac{1}{4}\right)\int_{\frac{1}{4}}^{\frac{1}{2}}P_n(t-x)dt.$$

It is easy to see that all the terms in x associated to $(j\circ f)(x_k)$ are positive and therefore we obtain the form

$$K_n(\overline{R}_n(j\circ f),x)=\sum_{k=-n}^{n}(j\circ f)(x_k)p_{n,k}(x)$$

with $p_{n.k}(x)\geq 0$ for all $x\in\left[-\frac{1}{4},\frac{1}{4}\right]$.

With the help of $p_{n.k}(x)$ given as above, we define the Szabados-type fuzzy generalized polynomial associated to $f:\left[-\frac{1}{4},\frac{1}{4}\right]\to\mathbb{R}_{\mathcal{F}}$, by

$$S(x)=\sum_{k=-n}^{n}p_{n,k}(x)\odot f\left(x_k\right).$$

The following theorem gives Jackson-type estimate for the error of approximation by Szabados-type polynomials.

Theorem 3.9. *Let* $f:\left[-\frac{1}{4},\frac{1}{4}\right]\to\mathbb{R}_{\mathcal{F}}$ *be continuous and*

$$S(x)=\sum_{k=-n}^{n}p_{n,k}(x)\odot f\left(x_k\right)$$

defined as above. Then

$$D^*(f,S)\leq C\omega_1^{\mathcal{F}}\left(f,\frac{1}{n}\right)_{\left[-\frac{1}{4},\frac{1}{4}\right]}.$$

Proof. Since all $p_{n,k}(x)\geq 0$, we have

$$\sum_{k=-n}^{n}(j\circ f)(x_k)p_{n,k}(x)=j\left(\sum_{k=-n}^{n}p_{n,k}(x)\odot f(x_k)\right).$$

But j is an isometry, so we have:

$$D(f(x),S(x))=D\left(f(x),\sum_{k=-n}^{n}p_{n,k}(x)\odot f(x_k)\right)$$

$$
\begin{aligned}
&= \left\| (j \circ f)(x) - j\left(\sum_{k=-n}^{n} p_{n,k}(x) \odot f(x_k) \right) \right\| \\
&= \left\| (j \circ f)(x) - \sum_{k=-n}^{n} (j \circ f)(x_k) p_{n,k}(x) \right\|,
\end{aligned}
$$

where $\|\cdot\|$ is the norm in $\mathbb{B} = \overline{C}[0,1] \times \overline{C}[0,1]$.

We observe that the last sum is $K_n(\overline{R}_n(j \circ f), x)$. Also we have:

$$
\left\| (j \circ f)(x) - K_n(\overline{R}_n(j \circ f), x) \right\| \leq \| (j \circ f)(x) - R_n(j \circ f, x) \| +
$$

$$
+ \left\| R_n(j \circ f, x) - K_n(\overline{R}_n(j \circ f), x) \right\|.
$$

Since the coefficients $r_{n,k}(x)$ of R_n in Lemma 3.4 are all positive, we have

$$
R_n(j \circ f, x) = \sum_{k=-n}^{n} r_{n,k}(x)(j \circ f)(x_k) = j\left(\sum_{k=-n}^{n} p_{n,k}(x) \odot f(x_k) \right)
$$

and taking into account that j is an isometry, we obtain for all $x \in \left[-\frac{1}{4}, \frac{1}{4}\right]$

$$
\begin{aligned}
D(f(x), S(x)) &\leq D(f(x), R_n(f, x)) + \left\| R_n(j \circ f, x) - K_n(\overline{R}_n(j \circ f), x) \right\| = \\
&= D(f(x), R_n(f, x)) + \left\| \overline{R}_n(j \circ f, x) - K_n(\overline{R}_n(j \circ f), x) \right\|.
\end{aligned}
$$

(this last inequality is obvious by the definition of $\overline{R}_n$).

By Lemma 3.4 and Lemma 3.8 we obtain:

$$
D(f(x), S(x)) \leq 5\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{\left[-\frac{1}{4}, \frac{1}{4}\right]} + C_4 \omega_1^{\mathbb{B}}\left(\overline{R}_n(j \circ f), \frac{1}{n}\right)_{\left[-\frac{1}{4}, \frac{1}{4}\right]}.
$$

It is easy to see that $\omega_1^{\mathbb{B}}\left(\overline{R}_n(j \circ f), \frac{1}{n}\right)_{\left[-\frac{1}{4}, \frac{1}{4}\right]} = \omega_1^{\mathbb{B}}\left(R_n(j \circ f), \frac{1}{n}\right)_{\left[-\frac{1}{4}, \frac{1}{4}\right]}$. By Lagrange theorem for functions with values in Banach spaces we obtain:

$$
\| R_n(j \circ f, y) - R_n(j \circ f, x) \| \leq \sup_{\xi \in [x,y]} \| R_n'(j \circ f, \xi) \|_{C(I', \mathbb{B})} \cdot |y - x|
$$

with $I' = \left[-\frac{1}{4}, \frac{1}{4}\right]$ and for $|y - x| \leq \frac{1}{n}$, taking into account Lemma 3.7 we obtain

$$
\omega_1^{\mathbb{B}}\left(\overline{R}_n(j \circ f), \frac{1}{n}\right)_{I'} \leq 900n \cdot \omega_1^{\mathbb{B}}\left(j \circ f, \frac{1}{n}\right)_{I'} \cdot \frac{1}{n} = 900\omega_1^{\mathbb{B}}\left(j \circ f, \frac{1}{n}\right)_{I'}.
$$

It is easy to check that $\omega_1^{\mathbb{B}}\left(j \circ f, \frac{1}{n}\right)_{I'} = \omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{I'}$ and we finally obtain

$$D(f(x), S(x)) \leq 5\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{I'} + C_4 \cdot 900\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{I'} = C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{I'}$$

which completes the proof. □

As an immediate consequence we obtain the following Jackson-type estimate for the error in approximation by generalized fuzzy algebraic polynomials.

Corollary 3.10. *For the best approximation by algebraic polynomials we have*

$$E_n^{K,1} \leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{\left[-\frac{1}{4}, \frac{1}{4}\right]}, \ \forall n \in \mathbb{N},\ f \in C_{\mathcal{F}}\left[-\frac{1}{4}, \frac{1}{4}\right],\ K = f\left(\left[-\frac{1}{4}, \frac{1}{4}\right]\right).$$

where $C > 0$ is an absolute constant independent of n and f.

Proof. Since the polynomial $P_n(t-x) \geq 0$, $\forall t, x \in [-1, 1]$ and $\int_{-1}^{1} P_n(t-x)dt = 1$, we have

$$|p_{n,-n}(x)| = \int_{-\frac{1}{2}}^{-\frac{1}{4}} P_n(t-x)dt + \int_{-\frac{1}{4}}^{\frac{1}{4}} \frac{P_n(t-x)dt}{(t-x_{-n})^4 \sum_{j=-n}^{n}(t-x_j)^{-4}} \leq$$
$$\leq \int_{-\frac{1}{2}}^{\frac{1}{4}} P_n(t-x)dt \leq 1.$$

Also

$$|p_{n,n}(x)| = \int_{\frac{1}{4}}^{\frac{1}{2}} P_n(t-x)dt + \int_{-\frac{1}{4}}^{\frac{1}{4}} \frac{P_n(t-x)dt}{(t-x_n)^4 \sum_{j=-n}^{n}(t-x_j)^{-4}} \leq 1.$$

and

$$|p_{n,0}(x)| = 1 - \int_{-\frac{1}{2}}^{\frac{1}{2}} P_n(t-x)dt + \int_{-\frac{1}{4}}^{\frac{1}{4}} \frac{P_n(t-x)dt}{(t-x_0)^4 \sum_{j=-n}^{n}(t-x_j)^{-4}} \leq$$
$$\leq 1 - \int_{-\frac{1}{2}}^{\frac{1}{2}} P_n(t-x)dt + \int_{-\frac{1}{4}}^{\frac{1}{4}} P_n(t-x)dt \leq 1.$$

For $k \notin \{-n, 0, n\}$ we have

$$|p_{n,k}(x)| = \int_{-\frac{1}{4}}^{\frac{1}{4}} \frac{P_n(t-x)dt}{(t-x_k)^4 \sum_{j=-n}^{n}(t-x_j)^{-4}} \leq \int_{-\frac{1}{4}}^{\frac{1}{4}} P_n(t-x)dt \leq 1.$$

and the proof is complete. □

Remark 3.11. We can obtain the above results in any interval $[a, b]$ instead of $\left[-\frac{1}{4}, \frac{1}{4}\right]$ by mapping this interval in $[a, b]$ through a linear fnction which maps $-\frac{1}{4}$ to a and $\frac{1}{4}$ to b.

4 Application to fuzzy interpolation

In this section we prove the convergence of fuzzy Lagrange polynomials for some classes of fuzzy functions.

The fuzzy Lagrange plynomial is defined in [14], [9] as follows (see also [8, p.651]) $L_n(x) = \sum_{i=0}^{n} l_i(x) \odot f(x_i)$, where $l_i(x) = \frac{(x-x_0).../...(x-x_n)}{(x_i-x_0).../...(x_i-x_n)}$ are the usual fundametal Lagrange interpolation polynomials and the sign "/" means that the $i^{\underline{th}}$ operand is missing.

Theorem 4.1. *Let $f : [-1, 1] \to \mathbb{R}_{\mathcal{F}}$ be a Lipschitz mapping of order $\alpha > \frac{1}{2}$(i. e. there exists L such that $D(f(x), f(y)) \leq L\,|x-y|^{\alpha}$ for all $x, y \in [-1, 1]$. Let $(x_{n,i})_{i=\overline{1,n}}$, $n \in \mathbb{N}$ be a normal matrix of nodes and $L_n(x)$ the fuzzy Lagrange polynomial which interpolates f on $\{x_{n,0}, ..., x_{n,n}\}$. Then*

$$\lim_{n\to\infty} L_n(x) = f(x),\ \forall x \in [-1, 1].$$

The convergence is uniform in any interval $[-1+h, 1-h]$, $0 < h < 1$.

Proof. By Corollary 3.10 if we take $M = \max\left\{1, \sqrt{\frac{n}{2h}}\right\}$, $\forall h > 0$, then $E_n^{K,M}(f) \leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{[-1,1]}$, where $K = f([-1, 1])$ and also the best approximation polynomial in $A_n^{K,M}$ (denoted π_n) exists. By [5, Lemma 8.3.2, p. 351] for a normal matrix of nodes we have $|l_i(x)| \leq \sum_{i=0}^{n} |l_i(x)| \leq \sqrt{\frac{n}{2h}}$ for $x \in [-1+h, 1-h]$ and so $L_n \in A_n^{K,M}$. Then $D^*(L_n, f) \leq D^*(L_n, \pi_n) + D^*(\pi_n, f)$. By Corollary 3.10 we obtain $D^*(\pi_n, f) \leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{[-1,1]}$.

Let $L_n(\pi_n)$ be the fuzzy Lagrange polynomial associated to π_n at $\{x_{n.0}, ..., x_{n,n}\}$. We prove that $L_n(\pi_n) = \pi_n$. We observe that $\pi_n(x_{n,j}) = \sum_{i=0}^{n} l_i(x) \odot \pi(x_{n,i})$ since $l_i(x_{n,j}) = \delta_{i,j}$ (Kronecker symbol $\delta_{i,j}$).

So $L_n(\pi_n)(x_{n,j}) = \pi_n(x_{n,j})$, $j = \overline{0,n}$. Since the Lagrange polynomial is unique (see [8, p.650]), we get $L_n(\pi_n) = \pi_n$. Then using the properties of the metric D we obtain:

$$D\left(L_n(x), \pi_n(x)\right) = D\left(\sum_{i=0}^{n} l_i(x) \odot f\left(x_{n,i}\right), \sum_{i=0}^{n} l_i(x) \odot \pi\left(x_{n,i}\right)\right) \leq$$

$$\leq \sum_{i=0}^{n} D(l_i(x) \odot f\left(x_{n,i}\right), l_i(x) \odot \pi\left(x_{n,i}\right)) \leq \sum_{i=0}^{n} |l_i(x)|\, D(f\left(x_{n,i}\right), \pi\left(x_{n,i}\right)).$$

Using again Corollary 3.10 we obtain

$$D\left(L_n(x), \pi_n(x)\right) \leq \sum_{i=0}^{n} |l_i(x)|\, C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{[-1,1]}.$$

By [5, Lemma 8.3.2, p. 351], for a normal matrix of nodes we have

$$\sum_{i=0}^{n} |l_i(x)| \leq \sqrt{\frac{(b-a)n}{h}}, \text{ for } x \in [a+h, b-h].$$

Then

$$D^*(L_n, f) \leq C\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{[-1,1]}\left(1 + \sqrt{\frac{2}{h}}\sqrt{n}\right).$$

Since f is of Lipschitz-type, we ave $\omega_1^{\mathcal{F}}\left(f, \frac{1}{n}\right)_{[-1,1]} \leq L\frac{1}{n^\alpha}$, $\alpha > \frac{1}{2}$. Then

$$D^*\left(L_n, f\right) \leq CL\frac{1}{n^\alpha} + CL\sqrt{\frac{2}{h}}\frac{1}{n^{\alpha-\frac{1}{2}}},$$

which competes the proof. □

References

[1] G. A. Anastassiou, Rate of convergence of fuzzy neural network operators, univariate case, J. Fuzzy Math. 10, No. 3(2002), 755-780.

[2] G. A. Anastassiou and S. G. Gal, Approximation Theory. Moduli of Continuity and Global Smoothness Preservation, Birkhäuser, Boston, Basel, Berlin, 2000.

[3] G. A. Anastassiou and S. G. Gal, On a fuzzy trigonometric approximation theorem of Weierstrass-type, J. Fuzzy Math., 9(2001), 47-56.

[4] B. Bede and S. G. Gal, Quadrature rules for integrals of fuzzy-number-valued functions, Fuzzy Sets and Systems, accepted.

[5] E.K. Blum, Numerical Analysis and Computation Theory and Practice, Addison-Wesley Series in Mathematics, Addison-Wesley Publishing Company, XII, 1972.

[6] P. Diamond, P. Kloeden, Metric Spaces of Fuzzy Sets, World Scientific, New Jersey, 1994.

[7] P. Diamond, P. Kloeden, A. Vladimirov, Spikes, broken planes and the approximation of convex sets, Fuzzy Sets and Systems 99(1999), 225-232.

[8] S.G. Gal, Approximation Theory in Fuzzy Setting, Chapter 13 in Handbook of Analytic Computational Methods in Applied Mathematics, (G. A. Anastassiou ed.) Chapman & Hall/CRC, Boca Raton, London, New York, Washington D.C., 2000.

[9] O. Kaleva, Interpolation of fuzzy data, Fuzzy Sets and Systems 61(1994), 63-70.

[10] Puyin Liu, Analysis of approximation of continuous fuzzy functions by multivariate fuzzy polynomials, Fuzzy Sets and Systems 127(2002), 299-313.

[11] W. Lodwick, J. Santos, Constructing consistent fuzzy surfaces from fuzzy data, Fuzzy Sets and Systems, 135(2003), 259-277.

[12] G.G. Lorentz, Appproximation of Functions, Chelsea Publishing Company, New York, 1986 (second edition).

[13] G.G. Lorentz and R. A. DeVore, Constructive Approximation, Polynomials and Splines Approximation, Springer-Verlag, New York, Berlin, Heidelberg 1993.

[14] R. Lowen, A fuzzy Lagrange interpolation theorem, Fuzzy Sets and Systems 34(1990) 33-38.

[15] J. Szabados, On a problem of R. DeVore, Acta Math. Hungar., 27 (1-2)(1976) 219-223.

[16] Wu Congxin and Gong Zengtai, On Henstock integral of fuzzy-number-valued functions, I, Fuzzy Sets and Systems, 115(2000), no. 3 , 377-391.

[17] Wu Congxin and Liu Danhong, A fuzzy Weierstrass approximation theorem, J. Fuzzy Math., 7 (1999), no. 1, 101-104.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,233-248,
2004,

Razumikhin Technique and Stability of Impulsive Differential- Difference Equations in Terms of Two Measures

Ivanka M. Stamova
Bourgas Free University, 8000 Bourgas, Bulgaria
E-mail: stamova@bfu.bg

Jahan Eftekhar
University of Texas at San Antonio,
San Antonio, Texas 78249-0665, U.S.A.

Abstract. The present paper deals with the investigation of the stability of the solutions of impulsive differential-difference equations with fixed moments of impulse effect. By means of differential inequalities on piecewise continuous functions coupled with the Razumikhin technique sufficient conditions for stability in terms of two different piecewise continuous measures of such equations are found. A possible application to the automatic control of vehicle height in active suspension systems is indicated.

2000 AMS Subject Classification 34K45

Key words : *Stability in terms of two measures*, *Razumikhin technique*, *impulsive differential − difference equations*, *active suspension*, *vehicle height control*

1. Introduction

The impulsive differential equations present an adequate mathematical model of many natural processes and phenomena investigated in science and engineering. Due to its diverse types of applications, the theory of impulsive equations has been developed very intensively for the last years[1-3, 12-14]. The

qualitative theory of functional differential equations is also well developed [8-11]. The impulsive differential-difference equations are generalization of impulsive differential equations (without delay) and of differential difference equations (without impulses). Their theory is analytically more attractive than the theory of impulsive ordinary differential equations[1, 4-7, 18].

In the present paper we study the stability of the solutions of impulsive systems of differential-difference equations with fixed moments of impulse effect in terms of two different piecewise continuous measures. The priorities of this approach are useful and well known in the investigations on the stability of the solutions of differential equations, as well as in the generalizations obtained by this method[8, 14, 15].

In order to study the stability of the solutions of impulsive systems under considerations, we use piecewise continuous auxiliary functions that are analogous to the classical Lyapunov functions. We are also concerned with differential inequalities on piecewise continuous functions. By means of this technique the investigation on the stability of solutions of impulsive differential-difference systems can be replaced by the studying of the stability of solutions of scalar impulsive differential equations. Our analysis employs some minimal subsets of a suitable space of piecewise continuous functions. The derivatives of the auxiliary piecewise continuous functions by the elements of these subsets are then estimated[4-7, 13, 16].

Such systems seem to have application, among other things, in the study of active suspension height control. In the interest of improving the overall performance of automotive vehicles, in recent years, suspension incorporating active components have been developed. The designs may cover a spectrum of of performance capabilities[17], but the active components alter only the vertical force reactions of the suspensions, not the kinematics. The conventional passive suspensions consist of usual components with spring and damping properties, which are time-invariant. The interest in active or semi-active suspensions derives from the potential for improvements to vehicle ride performance with no compromise or enhancement in handling. The full active suspensions incorporate actuators to generate the desired forces in the suspension. They actuators are normally hydraulic cylinders.

2. Preliminary Notes and Definitions

Let $R_+ = [0, \infty)$ and let R^n be the n-dimensional Euclidean space with norm $|.|$. Let $t_0 \in R$ and $r > 0$.

Consider the impulsive system of differential-difference equations:

$$\begin{cases} \dot{x}(t) = f(t, x(t), x(t-r)), \quad t > t_0, \ t \neq \tau_k, \\ \Delta x(\tau_k) = I_k(x(\tau_k)), \quad \tau_k > t_0, \ k = 1, 2, ..., \end{cases} \tag{1}$$

where $f : (t_0, \infty) \times R^n \times R^n \to R^n$; $I_k : R^n \to R^n$, $k = 1, 2, ...$; $\Delta x(\tau_k) = x(\tau_k + 0) - x(\tau_k - 0)$; $t_0 \equiv \tau_0 < \tau_k < \tau_{k+1} < ...$ and $\lim\limits_{k\to\infty} \tau_k = \infty$.

Let $\phi \in C[[t_0 - r, t_0], R^n]$. Denote by $x(t) = x(t; t_0, \phi)$ the solution of the system (1), that satisfies the initial conditions

$$x(t, t_0, \phi) = \phi(t), \ t \in [t_0 - r, t_0],$$

$$x(t_0 + 0, t_0, \phi) = \phi(t_0).$$

The solutions $x(t) = x(t; t_0, \phi)$ of systems of the type (1) are piecewise continuous functions with points of discontinuity of the first kind $\tau_k > t_0$, $k = 1, 2, ...$ at which they are continuous from the left, i.e., at the moments of impulse effect τ_k the following relations are valid

$$x(\tau_k - 0) = x(\tau_k),$$

$$x(\tau_k + 0) = x(\tau_k) + I_k(x(\tau_k)), \ k = 1, 2, ...$$

Together with the system (1), we consider the scalar impulsive differential equation

$$\begin{cases} \dot{u}(t) = g(t, u(t)), \quad t \neq \tau_k, \ k = 1, 2, ..., \\ \Delta u(\tau_k) = B_k(u(\tau_k)), \ k = 1, 2, ..., \end{cases} \tag{2}$$

where $g : [t_0, \infty) \times R_+ \to R$ and $B_k : R_+ \to R$, $k = 1, 2, ...$

We introduce the following notations

$$G_k = \{(t, x) \in [t_0, \infty) \times R^n : \ \tau_{k-1} < t < \tau_k\}, k = 1, 2, ...;$$

$$G = \bigcup_{k=1}^{\infty} G_k.$$

Definition 1. We shall say that the function $V : [t_0, \infty) \times R^n \to R_+$ belongs to the class V_0 if:

1. The function V is continuous in G and locally Lipschitz continuous with respect to its second argument in each of the sets G_k, $k = 1, 2, ...$

2. For each $k = 1, 2, ...$ and $x \in R^n$ there exist the finite limits

$$V(\tau_k - 0, x) = \lim_{\substack{t\to\tau_k \\ t<\tau_k}} V(t, x), \ V(\tau_k + 0, x) = \lim_{\substack{t\to\tau_k \\ t>\tau_k}} V(t, x).$$

3. The equality $V(\tau_k - 0, x) = V(\tau_k, x)$ is valid.

In the sequel we will use the next classes of functions:
$K = \{a \in C[R_+, R_+]\colon\ a(r)$ is strictly increasing and $a(0) = 0\ \}$;
$CK = \{a \in C[[t_0, \infty) \times R_+, R_+]\colon\ a(t, .) \in K$ for any fixed $t \in [t_0, \infty)\ \}$;
$PC[[t_0, \infty), R^n] = \{x : [t_0, \infty) \to R^n\colon\ x$ is piecewise continuous with points of discontinuity of the first kind $\tau_k,\ k = 1, 2, ...$ at which it is continuous from the left $\}$;
$C_0 = C[[t_0 - r, t_0], R^n]$; $\Gamma = \{h \in V_0\colon \inf\limits_{x \in R^n} h(t, x) = 0$ for each $t \in [t_0, \infty)\ \}$;
$\Gamma_0 = \{h_0 \in C[[t_0 - r, t_0] \times C_0, R_+]\colon \inf\limits_{\phi \in C_0} h(t, \phi) = 0$ for each $t \in [t_0 - r, t_0]\ \}$;
$\Omega_0 = \{x \in PC[[t_0, \infty), R^n] :\ V(s, x(s)) \leq L(V(t, x(t))),\ t - r < s \leq t, t \geq t_0\}$;
$\Omega_1 = \{x \in PC[[t_0, \infty), R^n] :\ V(s, x(s)) \leq V(t, x(t)),\ t - r < s \leq t, t \geq t_0\}$;
$\Omega_A = \{x \in PC[[t_0, \infty), R^n] : A(s)V(s, x(s)) \leq A(t)V(t, x(t)),\ t - r < s \leq t, t \geq t_0, \}$.

In the above notations: $A : [t_0, \infty) \to (0, \infty)$ is a piecewise continuous function, having points of discontinuity of the first kind τ_k, A is continuous from the left at τ_k, $A(\tau_k + 0) > 0$ and $A(t) \to \infty\ as\ t \to \infty$; $L(u)$ is continuous on R_+, nondecreasing in u, and $L(u) > u$ for $u > 0$.

Assume $\rho > 0,\ \ h \in \Gamma$ and let

$$S(h, \rho) = \{(t, x, y) \in [t_0, \infty) \times R^n \times R^n :\ h(t, x) < \rho,\ h(t, y) < \rho\};$$

$$B(h, \rho) = \{(t, x) \in [t_0, \infty) \times R^n :\ h(t, x) < \rho\}.$$

Introduce the following conditions :

A1. The function $f : S(h, \rho) \to R^n$ is continuous in $S(h, \rho)$.

A2. $I_k \in C[R^n, R^n],\ \ k = 1, 2, ...$

A3. $t_0 = \tau_0 < \tau_1 < \tau_2 < ...$ and $\lim\limits_{k \to \infty} \tau_k = \infty$.

A4. $g \in C[[t_0, \infty) \times R_+, R]$ and $g(t, 0) = 0$ for $t \in [t_0, \infty)$.

A5. $B_k \in C[R_+, R]$, $B_k(0) = 0$ and $\psi_k(u) = u + B_k(u)$ are nondecreasing with respect to u, $k = 1, 2, ...$

A6. There exists $\rho_0,\ \ 0 < \rho_0 < \rho$, such that $h(\tau_k, x) < \rho_0$ implies $h(\tau_k + 0, x + I_k(x)) < \rho,\ \ k = 1, 2, ..., h \in \Gamma$.

Definition 2. Let $h \in \Gamma,\ h_0 \in \Gamma_0$.

(a) h_0 is *finer* than h if there exist a number $\delta > 0$ and a function $\varphi \in K$ such that $h_0(t, \phi) < \delta$ leads to $h(t, x) \leq \varphi(h_0(t, \phi))$.

(b) h_0 is *weakly finer* than h if there exist a number $\delta > 0$ and a function $\varphi \in CK$ such that $h_0(t, \phi) < \delta$ leads to $h(t, x) \leq \varphi(t, h_0(t, \phi))$.

Let $V \in V_0,\ x \in PC[[t_0, \infty), R^n]$ and $\ t \neq \tau_k,\ k = 1, 2, ...$ Introduce the function

$$D_{-}V(t, x(t)) = \lim_{\sigma \to 0^-} inf \sigma^{-1}[V(t + \sigma, x(t) + \sigma f(t, x(t), x(t - r))) - V(t, x(t))].$$

We will use the following definitions of stability of the system (1) in terms of two different measures, that generalize various classical notions of stability.

Definition 3. For $h \in \Gamma,\ h_0 \in \Gamma_0$ the system (1) is said to be:

(a) (h_0, h) - *stable* if

$$(\forall t_0 \in R)(\forall \varepsilon > 0)(\exists \delta = \delta(t_0, \varepsilon) > 0)$$
$$(\forall \phi \in C_0 : \max_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s)) < \delta)$$
$$(\forall t \geq t_0) : \quad h(t, x(t; t_0, \phi) < \varepsilon.$$

(b) (h_0, h) - *uniformly stable* if the number δ from (a) does not depend on t_0.

(c) (h_0, h) - *equiattractive* if

$$(\forall t_0 \in R)(\exists \delta = \delta(t_0) > 0)(\forall \varepsilon > 0)(\exists T = T(t_0, \varepsilon) > 0)$$
$$(\forall \phi \in C_0 : \quad \max_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s)) < \delta)(\forall t > t_0 + T) : \quad h(t, x(t; t_0, \phi) < \varepsilon.$$

(d) (h_0, h) - *uniformly attractive* if the numbers δ and T from (c) are independent on t_0.

(e) (h_0, h) - *equiasymptotically stable* if it is (h_0, h) - stable and (h_0, h) - equiattractive.

(f) (h_0, h) - *uniformly asymptotically stable* if it is (h_0, h) - uniformly stable and (h_0, h) - uniformly attractive.

For a concrete choice of the measures h_0 and h Definition 3 is reduces to the following particular cases:

1) Lyapunov's stability of the zero solution of (1) if $h_0(t, \phi) = ||\phi|| = \max_{t \in [t_0 - r, t_0]} |\phi(t)|$ and $h(t, x) = |x|$.

2) stability by part of the variables of the zero solution of (1) if

$$h_0(t, \phi) = ||\phi||,\ \ h(t, x) = |x|_k = \sqrt{x_1^2 + ... + x_k^2},$$
$$x = (x_1, ..., x_n),\ \ 1 \leq k \leq n.$$

3) Lyapunov's stability of the non-null solution $x_0(t) = x_0(t;t_0,\phi_0)$ of (1) if $h_0(t,\phi) = ||\phi - \phi_0||$, $h(t,x) = |x - x_0(t)|$.

4) stability of the set $M \subset [t_0 - r, \infty) \times R^n$ if
$h_0(t,\phi) = \max\limits_{t\in[t_0-h,t_0]} d(\phi(t), M_0(t))$ and $h(t,x) = d(x, M(t))$, where

$$M(t) = \{x \in R^n : \ (t,x) \in M, \ t > t_0\},$$

$$M_0(t) = \{x \in R^n : \ (t,x) \in M, \ t \in [t_0 - h, t_0]\}.$$

5) stability of conditionally invariant set B with respect to the set A, where $A \subset B \subset R^n$ if

$$h_0(t,\phi) = d(\phi, A), \quad h(t,x) = d(x, B).$$

Definition 4. Let $h \in \Gamma$, $h_0 \in \Gamma_0$ and $V \in V_0$. The function V is said to be:

(a) h - *positively definite* if there exist a number $\delta > 0$ and a function $a \in K$ such that $h(t,x) < \delta$ implies $V(t,x) \geq a(h(t,x))$.

(b) h_0 - *decrescent* if there exist a number $\delta > 0$ and a function $b \in K$ such that $h_0(t,\phi) < \delta$ implies $V(t+0,x) \leq b(h_0(t,\phi))$.

(c)*weakly* h_0 - *decrescent* if there exist a number $\delta > 0$ and a function $b \in CK$ such that $h_0(t,\phi) < \delta$ implies $V(t+0,x) \leq b(t, h_0(t,\phi))$.

3. **Basic Comparison Theorems**

In the proofs of our main results we need the following comparison theorems.

Theorem 1. [6] *Assume the following conditions hold* :

1. *Assumptions* $A1 - A5$ *are valid.*

2. *The function* $V \in V_0$, $V : B(h,\rho) \to R_+$ *is such that for* $t \geq t_0$ *and* $x \in \Omega_1$ *we have*

$$\begin{cases} D_{-}V(t,x(t)) \leq g(t, V(t,x(t))), \ t \neq \tau_k, \ k = 1,2,..., \\ V(\tau_k + 0, x(\tau_k) + I_k(x(\tau_k))) \leq \psi_k(V(\tau_k, x(\tau_k))), \ k = 1,2,.... \end{cases}$$

3. *For the solution* $x(t;t_0,\phi)$ *of the system* (1) *we have* $(t, x(t+0;t_0,\phi)) \in B(h,\rho)$ *as* $t \in [t_0,\infty)$ *and* $h \in \Gamma$.

4. *The maximal solution* $r(t;t_0,u_0)$, $u_0 \geq V(t_0+0, \phi(t_0))$, *of the equation* (2) *is defined on the interval* $[t_0,\infty)$.

Then

$$V(t, x(t;t_0,\phi)) \leq r(t;t_0,u_0) \ \ for\ t \in [t_0,\infty).$$

Corollary 1. *Let the following conditions hold* :

1. *Assumptions* A1-A3 *are met.*

2. *The function* $V \in V_0,\ V : B(h, \rho) \to R_+$ *is such that for* $t \geq t_0$ *and* $x \in \Omega_0$ *we have*

$$D_{-}V(t, x(t)) \leq 0,\ t \neq \tau_k,\ k = 1, 2, ...,$$

$$V(\tau_k + 0, x(\tau_k) + I_k(x(\tau_k))) \leq V(\tau_k, x(\tau_k)),\ \ k = 1, 2, ...$$

3. *Condition* 3 *of Theorem* 1 *holds.*

Then

$$V(t, x(t; t_0, \phi)) \leq V(t_0 + 0, \phi(t_0)),\ t \in [t_0, \infty).$$

Theorem 2. *Assume the following conditions hold* :

1. *Assumptions* $A1 - A5$ *are valid.*

2. *The function* $V \in V_0,\ V : B(h, \rho) \to R_+$ *is such that for* $t \geq t_0$ *and* $x \in \Omega_A$ *we have*

$$A(t)D_{-}V(t, x(t)) + V(t, x(t))D_{-}A(t)$$

$$\leq g(t, A(t)V(t, x(t))),\ t \neq \tau_k,\ k = 1, 2, ..., \tag{3}$$

$$A(\tau_k + 0)V(\tau_k + 0, x(\tau_k) + I_k(x(\tau_k))) \leq \psi_k(A(\tau_k)V(\tau_k, x(\tau_k))),\ k = 1, 2, ..., \tag{4}$$

where $A : [t_0, \infty) \to (0, \infty)$ *is a piecewise continuous function, having points of discontinuity* $\tau_k, k = 1, 2, ...$ *of first kind at which it is continuous from the left,* $A(\tau_k + 0) > 0,\ \ k = 1, 2, ...$ *and*

$$D_{-}A(t) = \lim_{\sigma \to 0^-} inf \sigma^{-1}[A(t + \sigma) - A(t)].$$

3. *Condition* 3 *of Theorem* 1 *holds.*

4. *The maximal solution* $r(t; t_0, u_0),\ u_0 \geq A(t_0 + 0)V(t_0 + 0, \phi(t_0))$ *of the equation* (2) *is defined on the interval* $[t_0, \infty)$.

Then

$$A(t)V(t, x(t; t_0, \phi)) \leq r(t; t_0, u_0)\ \ for\ t \in [t_0, \infty). \tag{5}$$

Proof. Setting

$$W(t, x(t)) = A(t)V(t, x(t))$$

and let $t \geq t_0$ and $x \in \Omega_A$. For $t \neq \tau_k,\ k = 1, 2, ...,$ and for sufficiently close to zero $\sigma < 0$ we have

$$W(t + \sigma, x(t) + \sigma f(t, x(t), x(t - r))) - W(t, x(t))$$

$$= V(t+\sigma, x(t)+\sigma f(t,x(t),x(t-r)))[A(t+\sigma)-A(t)]$$
$$+A(t)[V(t+\sigma, x(t)+\sigma f(t,x(t),x(t-r))) - V(t,x(t))].$$

It follows from (3) and (4) that

$$D_{-}W(t,x(t)) \le g(t, W(t,x(t))),\ t \ne \tau_k,\ k=1,2,...,$$
$$W(\tau_k+0, x(\tau_k)+I_k(x(\tau_k))) \le \psi_k(W(\tau_k, x(\tau_k))),\ k=1,2,....$$

Now the inequality (5) follows after application of Theorem 1 for the function $W(t,x)$.

4. Main Results

Theorem 3. *Assume the following conditions hold :*

1. *Condition* (A) *is valid.*
2. $h \in \Gamma,\ h_0 \in \Gamma_0$ *and* h_0 *is weakly finer than* h.
3. *The function* $V \in V_0,\ V : B(h,\rho) \to R_+$ *is* $h-$ *positively definite on* $B(h,\rho)$ *and it is weakly* h_0- *decrescent.*
4. *For* $t \ge t_0$ *and* $x \in \Omega_1$ *we have*

$$D_{-}V(t,x(t)) \le g(t, V(t,x(t))),\ t \ne \tau_k,\ k=1,2,...,$$
$$V(\tau_k+0, x(\tau_k)+I_k(x(\tau_k))) \le \psi_k(V(\tau_k, x(\tau_k))),\ k=1,2,....$$

Then

(a) if the zero solution of the equation (2) *is stable then the system* (1) *is* $(h_0,h)-$ *stable.*

(b) if the zero solution of the equation (2) *is equiasymptotically stable then the system* (1) *is* $(h_0,h)-$ *equiasymptotically stable.*

Proof. (a) Since V is h- positively definite on $B(h,\rho)$ then there exists a function $a \in K$ such that:

$$V(t,x) \ge a(h(t,x)) \quad as \quad (t,x) \in B(h,\rho). \tag{6}$$

On the other hand V is weakly h_0- decrescent and there exist a number $\delta_1 > 0$ and a function $b \in CK$ such that

$$V(t+0,x) \le b(t, h_0(t,\phi)) \quad as \quad h_0(t,\phi) < \delta_1. \tag{7}$$

By means of the second condition of Theorem 3 there exist $\delta_2 > 0$ and $\varphi \in CK$ such that

$$h(t,x) \le \varphi(t, h_0(t,\phi)) \quad as \quad h_0(t,\phi) < \delta_2.$$

Let $0 < \varepsilon < \rho_0$ and $t_0 \in R$. It follows from the properties of the function φ that there exist number $\delta_3 = \delta_3(t_0, \varepsilon)$, $0 < \delta_3 < \delta_2$ such that

$$\varphi(t_0, \delta_3) < \rho. \tag{8}$$

Now, the stability of the zero solution of (2) ensures that there exists $\delta_4 = \delta_4(t_0, \varepsilon)$ such that

$$r(t; t_0, u_0) < a(\varepsilon) \quad as \quad 0 \leq u_0 < \delta_4, \ t \geq t_0, \tag{9}$$

where $r(t; t_0, u_0)$ is the maximal solution of (2) satisfying $r(t_0 + 0; t_0, u_0) = u_0$.

We choose now the number $\delta_5 = \delta_5(t_0, \varepsilon)$ so that

$$b(t_0, \delta_5) < \delta_4. \tag{10}$$

Setting $\delta = \min(\delta_3, \delta_4, \delta_5)$ and let $\phi \in C_0$, $\max\limits_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s)) < \delta$ and $h_0(t_0 + 0, \phi(t_0)) < \delta$. It follows from (7) and (10) that

$$\begin{aligned} V(t_0 + 0, \phi(t_0)) &\leq b(t_0, h_0(t_0, \phi(t_0))) \leq b(t_0, \max_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s))) \\ &\leq b(t_0, \delta) \leq b(t_0, \delta_5) < \delta_4. \end{aligned} \tag{11}$$

Supposing now $x(t) = x(t; t_0, \phi)$ to be such solution of the system (1) that $\max\limits_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s)) < \delta$.

We will prove now that

$$h(t, x(t)) < \varepsilon \quad as \quad t \geq t_0.$$

Supposing the opposite, there exists $t^* > t_0$ such that $\tau_k < t^* \leq \tau_{k+1}$ for some fixed integer k and

$$h(t^*, x(t^*)) \geq \varepsilon \quad and \quad h(t, x(t)) < \varepsilon, \quad t_0 < t \leq \tau_k.$$

Since $0 < \varepsilon < \rho_0$, condition A6 shows that

$$h(\tau_k + 0, x(\tau_k + 0)) = h(\tau_k + 0, x(\tau_k) + I_k(x(\tau_k))) < \rho.$$

Therefore there exists t^0, $\tau_k < t^0 \leq t^*$, such that

$$\varepsilon \leq h(t^0, x(t^0)) < \rho \quad and \quad h(t, x(t)) < \rho, \quad t_0 < t \leq t^0. \tag{12}$$

Applying now Theorem 1 for the interval $(t_0, t^0]$ and $u_0 = V(t_0 + 0, \phi(t_0))$ we obtain

$$V(t, x(t; t_0, \phi)) \leq r(t; t_0, V(t_0 + 0, \phi(t_0))), \quad t_0 < t \leq t^0. \tag{13}$$

So the implications (12), (6), (13), (8) and (9) lead to

$$a(\varepsilon) \leq a(h(t^0, x(t^0))) \leq V(t^0, x(t^0))$$

$$\leq r(t^0; t_0, V(t_0 + 0, \phi(t_0))) < a(\varepsilon).$$

The contradiction we have already obtained shows that $h(t, x(t)) < \varepsilon$ for each $t \geq t_0$. Therefore the system (1) is (h_0, h)- stable.

(b) It follows from (a) that the system (1) is (h_0, h)- stable. So, for each $t_0 \in R$ there exists a number $\delta_{01} = \delta_{01}(t_0, \rho) > 0$ such that if $\phi \in C_0$, $\max\limits_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s)) < \delta_{01}$ then $h(t, x(t; t_0, \phi)) < \rho$ as $t \geq t_0$.

Let $0 < \varepsilon < \rho_0$ and $t_0 \in R$. The equiasymptotical stability of the zero solution of the equation (2) implies that there exist $\delta_{02} = \delta_{02}(t_0) > 0$ and $T = T(t_0, \varepsilon) > 0$ such that for $0 < u_0 < \delta_{02}$ and $t > t_0 + T$ the next inequality holds:

$$r(t; t_0, u_0) < a(\varepsilon). \tag{13}$$

We denote $\delta_{03} = \delta_{03}(t_0)$, $0 < \delta_{03} < \delta_{02}$ such that

$$b(t_0, \delta_{03}) < \delta_{02}. \tag{14}$$

It follows from (7) and (14) that if $\max\limits_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s)) < \delta_{03}$ then

$$V(t_0 + 0, \phi(t_0)) < b(t_0, h_0(t_0, \phi(0))) \leq b(t_0, \delta_{03}) < \delta_{02}.$$

In the case, by means of (13) we would have

$$r(t; t_0, V(t_0 + 0, \phi(t_0))) < a(\varepsilon), \quad t > t_0 + T. \tag{15}$$

Assume $\delta_0 = \min(\delta_{01}, \delta_{02}, \delta_{03})$ and let $\max\limits_{t_0 - r \leq s \leq t_0} h_0(s, \phi(s)) < \delta_0$. Theorem 1 shows that if $x(t) = x(t; t_0, \phi)$ is an arbitrary solution of the system (1) then

$$V(t, x(t; t_0, \phi)) \leq r(t; t_0, V(t_0 + 0, \phi(t_0))), \quad t > t_0. \tag{16}$$

Therefore we obtain from (6), (15) and (16) that the inequalities

$$a(h(t, x(t))) \leq V(t, x(t)) \leq r(t; t_0, V(t_0 + 0, \phi(t_0))) < a(\varepsilon)$$

hold for each $t > t_0 + T$. Hence $h(t, x(t)) < \varepsilon$ as $t > t_0 + T$ which shows that the system (1) is (h_0, h)- equiattractive. This proves Theorem 3.

Theorem 4. *Assume the following conditions hold :*

1. *Assumption* (A) *holds.*

2. $h \in \Gamma,\ h_0 \in \Gamma_0$ *and* h_0 *is finer than* h.

3. *The function* $V \in V_0,\ V : B(h,\rho) \to R_+$ *is* $h-$ *positively definite on* $B(h,\rho)$ *and* h_0- *decrescent.*

4. *Condition* 4 *of Theorem* 3 *is valid.*

Then

(*a*) *if the zero solution of the equation* (2) *is uniformly stable then the system* (1) *is* $(h_0, h)-$ *uniformly stable.*

(*b*)*if the zero solution of the equation* (2) *is uniformly asymptotically stable then the system* (1) *is* $(h_0, h)-$ *uniformly asymptotically stable.*

The proof of Theorem 4 is analogous to the proof of Theorem 3 and we omit it. Let us note that in this case the numbers δ, δ_0 and T can be chosen independently of t_0.

Theorem 5. *Assume the following conditions hold :*

1. *Assumptions* A1-A3 *and* A6 *are met.*

2. *Conditions* 2 *and* 3 *of Theorem* 4 *are valid.*

3. *For each* $t \geq t_0$ *and* $x \in \Omega_0$ *we have*

$$D_-V(t, x(t)) \leq 0,\ t \neq \tau_k,\ k = 1, 2, ...,$$

$$V(\tau_k + 0, x(\tau_k) + I_k(x(\tau_k))) \leq V(\tau_k, x(\tau_k)),\ \ k = 1, 2, ...$$

Then the system (1) *is* $(h_0, h)-$ *uniformly stable.*

The proof of Theorem 5 could be done in the same way as in Theorem 3(a), using Corollary 1 now.

Theorem 6. *Assume the following conditions hold :*

1. *Conditions* 1-3 *of Theorem* 3 *are valid.*

2. *For each* $t \geq t_0$ *and* $x \in \Omega_A$ *we have*

$$A(t)D_-V(t, x(t)) + V(t, x(t))D_-A(t)$$

$$\leq g(t, A(t)V(t, x(t))),\ t \neq \tau_k,\ k = 1, 2, ...,$$

$$A(\tau_k + 0)V(\tau_k + 0, x(\tau_k) + I_k(x(\tau_k))) \leq \psi_k(A(\tau_k)V(\tau_k, x(\tau_k))),\ k = 1, 2, ...,$$

where $A : [t_0, \infty) \to (0, \infty)$ *is a piecewise continuous function, having points of discontinuity* $\tau_k, k = 1, 2, ...$ *of first kind at which it is continuous from the left,* $A(\tau_k + 0) > 0,\ \ k = 1, 2, ...$ *and* $A(t) \to \infty$ *as* $t \to \infty$.

Then, if the zero solution of the equation (2) *is stable then the system* (1) *is* $(h_0, h)-$ *equiasymptotically stable.*

Proof. Let $\lambda = \inf\limits_{t\in[t_0,\infty)} A(t)$. The properties of A mean that $\lambda > 0$.

Since the function V is h- positively definite on $B(h,\rho)$, then there exists a function $a \in K$ such that:

$$V(t,x) \geq a(h(t,x)) \quad as \quad (t,x) \in B(h,\rho). \tag{17}$$

Moreover V is weakly h_0- decrescent and there exist a number $\delta_1 > 0$ and a function $b \in CK$ such that

$$V(t+0,x) \leq b(t,h_0(t,\phi)) \quad as \quad h_0(t,\phi) < \delta_1. \tag{18}$$

Let $0 < \varepsilon < \rho_0$ and $t_0 \in R$. We set $\varepsilon_1 = \lambda a(\varepsilon)$.

The stability of the zero solution of the equation (2) implies that there exists $\delta^* = \delta^*(t_0, \varepsilon_1) > 0$ such that if $0 < u_0 < \delta^*$, then $r(t;t_0,u_0) < \varepsilon_1, \quad t > t_0$ where $r(t;t_0,u_0)$ is the maximal solution of (2) satisfying $r(t_0+0;t_0,u_0) = u_0$. Repeating the proof of Theorem 3(a) and replacing $a(\varepsilon)$ with ε_1, $V(t_0+0,\phi(t_0))$ with $A(t_0+0)V(t_0+0,\phi(t_0))$ we obtain that the system (1) is (h_0,h)- stable.

Therefore there exists $\delta_0 = \delta_0(t_0,\rho) > 0$ such that if $\max\limits_{t_0-r\leq s\leq t_0} h_0(s,\phi(s)) < \delta_0$ then $h(t,x(t;t_0,\phi)) < \rho$ as $t > t_0$.

Let $\delta_1 = \delta_1(t_0,\varepsilon) > 0$ be such that if $0 < u_0 < \delta_1$, then $r(t;t_0,u_0) < \varepsilon$ as $t > t_0$. We can suppose that δ_1 is continuous and strictly increasing function with respect to ε for a fixed t_0.

Now we choose the number ε such that

$$A(t_0+0)b(t_0,\delta_0) = \delta_1(t_0,\varepsilon). \tag{19}$$

Let $x(t) = x(t;t_0,\phi)$ be such solution of the system (1) that $\max\limits_{t_0-r\leq s\leq t_0} h_0(s,\phi(s)) < \delta_0$. It follows from (18) and (19) that

$$A(t_0+0)V(t_0+0,\phi(t_0)) \leq A(t_0+0)b(t_0,h_0(t_0,\phi(t_0)))$$
$$< A(t_0+0)b(t_0,\delta_0) = \delta_1$$

whence

$$r(t;t_0,A(t_0+0)V(t_0+0,\phi(t_0))) < \varepsilon. \tag{20}$$

On the other hand Theorem 2 yields

$$A(t)V(t,x(t)) \leq r(t;t_0,A(t_0+0)V(t_0+0,\phi(t_0))) \tag{21}$$

for all $t > t_0$.

Now (17), (20) and (21) imply

$$A(t)a(h(t,x(t))) \leq A(t)V(t,x(t))$$

$$\leq r(t;t_0, A(t_0+0)V(t_0+0,\phi(t_0))) < \varepsilon.$$

Therefore $h(t,x(t)) < a^{-1}(\varepsilon/A(t))$. Since $A(t) \to \infty$ as $t \to \infty$ it follows that there exists $T^* = T^*(t_0,\varepsilon) > 0$ such that $h(t,x(t)) < \varepsilon$ as $t > T^*$.

Setting $T = T(t_0,\varepsilon) = T^*(t_0,\varepsilon) - t_0$ we get

$$h(t,x(t)) < \varepsilon \quad as \quad t \geq t_0 + T,$$

that proves the (h_0,h)- equiattractivity of the system (1).

5. An Example

Let us consider the linear impulsive differential-difference equation

$$\begin{cases} \dot{x}(t) = -ax(t) + bx(t-r), \quad t \neq \tau_k, \\ \Delta x(\tau_k) = -\alpha_k x(\tau_k), \quad k = 1,2,..., \end{cases} \tag{22}$$

where $a,\ b,\ r > 0;\ \ 0 \leq \alpha_k \leq 2,\ k = 1,2,...;\ 0 < \tau_1 < \tau_2 < ...$ and $\lim\limits_{k\to\infty} \tau_k = \infty$.

Let $\phi \in C[[-r,0],R]$. Denote by $x(t) = x(t;t_0,\phi)$ the solution of the equation (22), for which

$$x(t,0,\phi) = \phi(t),\ t \in [-r,0].$$

Let $h_0(t,\phi) = ||\phi|| = \max\limits_{t\in[-r,0]} |\phi(t)|$ and $h(t,x) = |x|$. We consider the functions $A(t) = e^{\alpha t},\ \alpha > 0$ and $V(t,x) = x^2$. The sets Ω_1 and Ω_A are defined by

$$\Omega_1 = \Big\{x \in PC[R_+,R] :\ x^2(s) \leq x^2(t),\ t-r < s \leq t\Big\}$$

and

$$\Omega_A = \Big\{x \in PC[R_+,R] :\ e^{\alpha s}x^2(s) \leq e^{\alpha t}x^2(t),\ t-r < s \leq t\Big\}.$$

If $t \geq 0$ and $x \in \Omega_1$ we have

$$D_-V(t,x(t)) = -2ax^2(t) + 2bx(t)x(t-r)$$

$$\leq 2V(t,x(t))[-a+b], \quad t \neq \tau_k,\ k = 1,2,...$$

although $t \geq 0$ and $x \in \Omega_A$ imply

$$A(t)D_-V(t,x(t)) + V(t,x(t))D_-A(t)$$
$$= 2e^{\alpha t}x(t)[-ax(t) + bx(t-r)] + \alpha e^{\alpha t}x^2(t)$$
$$\leq [-2a + 2b + \alpha]e^{\alpha t}V(t,x(t)), \quad t \neq \tau_k,\ k = 1,2,....$$

Moreover

$$V(\tau_k + 0, x(\tau_k) - \alpha_k x(\tau_k)) = (1-\alpha_k)^2 V(\tau_k, x(\tau_k))$$
$$\leq V(\tau_k, x(\tau_k)),\ k = 1,2,..., \quad x \in \Omega_1,$$

$$A(\tau_k + 0)V(\tau_k + 0, x(\tau_k) - \alpha_k x(\tau_k)) = (1-\alpha_k)^2 A(\tau_k)V(\tau_k, x(\tau_k))$$
$$\leq A(\tau_k)V(\tau_k, x(\tau_k)),\ k = 1,2,..., \quad x \in \Omega_A.$$

Assume the inequality $a \geq b$ holds. Then Theorem 4(a) with $g(t,u) = 0$ and $B_k(u) = 0,\ \ k = 1,2,...$, shows that the zero solution of the equation (22) is uniformly stable in a Lyapunov sense.

Let the inequality $b \leq a - \varepsilon$ hold for some positive ε. Applying Theorem 4(b), we obtain that the zero solution of (22) is uniformly asymptotically stable.

If $b < \frac{2a-\alpha}{2}$ then the conditions of Theorem 6 are fulfilled as $g(t,u) = 0$ and $B_k(u) = 0,\ k = 1,2,....$ It follows that the zero solution of (22) is equiasymptotically stable.

References

[1] D. D. Bainov, E. Minchev and I. M. Stamova, Present state of the stability theory for impulsive differential equations, *Communications in Applied Analysis*, 2, 197-226 (1998).

[2] D. D. Bainov, P. S. Simeonov, *Systems with Impulse Effect: Stability, Theory and Applications,* Ellis Horwood, Chichester, 1989.

[3] D. D. Bainov, P. S. Simeonov, *Theory of Impulsive Differential Equations: Periodic Solutions and Applications,* Longman, Harlow, 1993.

[4] D. D. Bainov, I. M. Stamova, Second method of Lyapunov and existence of periodic solutions of linear impulsive differential-difference equations, *PanAmerican Mathematical Journal*, 7, 27-35 (1997).

[5] D. D. Bainov, I. M. Stamova, Second method of Lyapunov and comparison principle for impulsive differential-difference Equations, *J. Austral. Math. Soc. Ser. B*, 38, 489-505 (1997).

[6] D. D. Bainov, I. M. Stamova, Stability of sets for impulsive differential-difference equations with variable impulsive perturbations, *Communications on Applied Nonlinear Analysis*, 5, 69-81 (1998).

[7] D. D. Bainov, I. M. Stamova, Stability of the solutions of impulsive functional differential equations by Lyapunov's direct method, *The ANZIAM Journal*, 2, 269-278 (2001).

[8] X. Fu, L. Zhang, Razumikhin-type theorems on boundedness in terms of two measures for functional differential systems, *Dynamics Systems and Applications*, 6, 589-598 (1997).

[9] J. K. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York, Heidelberg, Berlin, 1977.

[10] J. K. Hale, V. Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag, 1993.

[11] V. B. Kolmanovskii, V. R. Nosov, *Stability of Functional Differential Equations*, Academic Press, 1986.

[12] V. Lakshmikantham, D. D. Bainov and P. S. Simeonov, *Theory of Impulsive Differential Equations*, World Scientific, Singapore, New Jersey, London, 1989.

[13] V. Lakshmikantham, S. Leela and A. A. Martynyuk, *Stability Analysis of Nonlinear Systems*, Marcel Dekker Inc., New York, 1989.

[14] V. Lakshmikantham, X. Liu, *Stability Analysis in Terms of Two Measures*, World Scientific, Singapore, 1993.

[15] X. Liu, S. Sivasundaram, On the direct method of Lyapunov in terms of two measures, J. Math. Phy. Sci., 26, 381-400 (1992.

[16] B. S. Razumikhin, *Stability of Systems with Retardation*, Nauka, Moscow, 1988 (in Russian).

[17] R. S. Sharp, D. A. Crolla, Road vehicle suspension system Design-A Review, *Vehicle Systems Dynamics*, 16, No.3, 167-192(1998).

[18] I. M. Stamova, G. T. Stamov, Lyapunov-Razumikhin method for impulsive functional differential equations and applications to the population dynamics, *Journal of Computational and Applied Mathematics*, 130, 163-171 (2001).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,249-256,
2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Asymptotic stability of solutions of a system for heat propagation with second sound

Salim A. Messaoudi
Mathematical Sciences Department,
KFUPM, Dhahran 31261,
Saudi Arabia.
E-mail : messaoud@kfupm.edu.sa

Tel : 00 966 3 860 4570
Fax : 00 966 3 860 2340

Abstract

In this work we establish an exponential decay result for the solutions of a certain initial-boundary value problem for a nonlinear hyperbolic system describing heat propagation by second sound.
Keywords heat, second sound, nonlinear, hyperbolic, exponential decay.
AMS Subject Classification : 35L45 - 35K05 - 35K65.

1 Introduction

In the absence of deformation and external sources, heat propagation in one spatial dimension body is governed by the following equation of balance of energy

$$e_t + q_x = 0, \tag{1.1}$$

where the internal energy e and the heat flux q are functions of (x, t) and a subscript denotes a partial derivative with respect to the relevant variable. In Fourier's theory of heat conduction, the internal energy depends on the absolute temperature only; i.e.

$$e = \hat{e}(\theta) \tag{1.2}$$

whereas the heat flux is given by the relation

$$q = -\kappa(\theta)\theta_x. \tag{1.3}$$

As a consequence, the system governing the evolution of the heat flux and the absolute temperature takes the form

$$\begin{aligned} q + \kappa(\theta)\theta_x = 0 \\ q_x + \hat{e}'(\theta)\theta_t = 0, \end{aligned}$$

where κ and $\hat{e}'$ are strictly positive functions characterizing the material in consideration. In the case when $\hat{e}'$ and κ are independent of θ we get the familiar linear heat equation

$$\theta_t = k\theta_{xx}, \quad k = \kappa/\hat{e}'.$$

This equation provides a useful description of heat conduction under a large range of conditions and predicts an infinite speed of propagation; that is, any thermal disturbance at one point has an instantaneous effect elsewhere in the body. This is not always the case. In fact, experiments showed that heat conduction in some dielectric crystals at low temperatures is free of this paradox (infinite speed propagation) and disturbances which are almost entirely thermal may propagate in a finite speed. This phenomenon in dielectric crystals is called second sound.

These observations go back to 1948, when Cattaneo [1] proposed, instead of Fourier's law, a new constitutive relation

$$\tau(\theta)q_t + q = -\kappa(\theta)\theta_x, \tag{1.4}$$

where τ and κ are strictly positive functions depending on the absolute temperature. Coleman, Fabrizio, and Owen [4] showed in 1982 that, if (1.4) is adopted then compatibility with thermodynamics requires that the internal energy given by (1.3) be modified. To derive the appropriate form of the internal energy we define the free energy function by

$$\psi = e - \theta\eta \tag{1.5}$$

where η is the entropy which satisfies, in the absence of external heat supply, the growth relation

$$\eta_t \geq -\left(\frac{q}{\theta}\right)_x. \tag{1.6}$$

We then combine (1.1), (1.5), and (1.6) to produce the Clausius-Duhem inequality

$$\psi_t + \theta\eta_t + \frac{q\theta_x}{\theta} \leq 0. \tag{1.7}$$

By using a theorem by Coleman and Mizel [2], one can show that a necessary and sufficient condition for (1.7) to hold is that

$$\begin{aligned} \psi &= \psi(\theta, q) = \psi_0(\theta) + \frac{1}{2}b(\theta)q^2 \\ \eta &= \eta(\theta, q) = -\psi_\theta(\theta, q), \end{aligned} \tag{1.8}$$

where

$$b(\theta) = \frac{\tau(\theta)}{\theta\kappa(\theta)}.$$

The sufficiency of (1.8) is clear by virtue of

$$\psi_t + \eta\theta_t = b(\theta)qq_t, \tag{1.9}$$

which, when combined with (1.4), yields

$$\psi_t + \eta\theta_t + \frac{q\theta_x}{\theta} = -\frac{q^2}{\kappa\theta} \leq 0 \tag{1.10}$$

and when combined with (1.1) gives

$$\eta\theta_t = -q_x - b(\theta)qq_t. \tag{1.11}$$

It also follows from (1.5) and (1.8) that the internal energy has the form

$$e = \tilde{e}(\theta, q) = e_0(\theta) + a(\theta)q^2, \tag{1.12}$$

where a is a function determined by τ and κ; in particular $a(\theta) > 0$.

It is physically reasonable to assume that there exists a temperature $\bar{\theta} > 0$ such that if $q = 0$ and $\theta = \bar{\theta}$ then the specific heat, the termal relaxation time, and the thermal conductivity are positive; i.e.,

$$e_0'(\bar{\theta}) > 0, \qquad \tau(\bar{\theta}) > 0, \qquad \kappa(\bar{\theta}) > 0. \tag{1.13}$$

By substituting in (1.1), the expression of e in (1.12), the system governing the evolution of θ and q becomes

$$\begin{gathered} q_x + (e_0'(\theta) + a'(\theta)q^2)\theta_t + 2a(\theta)qq_t = 0 \\ \tau(\theta)q_t + q + \kappa(\theta)\theta_x = 0. \end{gathered} \tag{1.14}$$

So by virtue of (1.13), this system is strictly hyperbolic in a neighborhood $\mathcal{V}$ - which can be taken as small as necessary - of the equilibrium state $(\bar{\theta}, 0)$.

Global existence and decay of classical solutions to the Cauchy problem, as well as to some initial boundary value problems, have been established by Coleman, Hrusa, and Owen [3]. They also showed that (θ, q) tends to the equilibrium state, however no rate of decay has been discussed. Concerning formation of singularities, Messaoudi [5] studied the following system

$$\begin{gathered} \tau(\theta)q_t + q + \kappa(\theta)\theta_x = 0 \\ e_0'(\theta)\theta_t + q_x = 0 \end{gathered}$$

and showed, under the same restrictions on τ, e_0', and κ, that classical solutions of the Cauchy problem break down in finite time if the initial data are chosen small in the L^∞ norm with large enough derivatives. This result has been later generalized by the author [6] to a system of the form

$$\begin{aligned} \sigma(e,q)q_t + \mu(e,q)q &= -e_x + \lambda(e,q)qq_x \\ e_t &= -q_x, \end{aligned} \tag{1.15}$$

where σ, μ satisfy

$$\sigma(\xi,\zeta) \geq \underline{\sigma} > 0, \quad \mu(\xi,\zeta) \geq \underline{\mu} > 0, \quad \forall(\xi,\zeta) \in \mathbb{R}^2. \tag{1.16}$$

We should note here that (1.15) is equivalent to (1.14). For more details, we refer the reader to [6].

In this article we consider system (1.14) together with the initial-boundary conditions

$$\begin{aligned} \theta(x,0) &= \theta_0(x), \quad q(x,0) = q_0(x), \quad x \in I = (0,1) && (1.17)\\ \theta(0,t) - \bar{\theta} &= \theta(1,t) - \bar{\theta} = 0, \quad t \geq 0 && (1.18) \end{aligned}$$

and establish an exponential decay result. In proving this result we make a crucial use of Poincare 's inequality to establish some estimates. This of course made our argument unextendable to unbounded intervals.

2 Exponential Decay

In this section we state and prove our main result.

Theorem. *Assume that e_0', a, κ, τ are C^2 functions satisfying (1.13). Then there exists a small positive constant δ such that for any initial data $\theta_0 - \bar{\theta} \in H^2(I) \cap H_0^1(I)$, $q_0 \in H^2(I)$ satisfying*

$$\|\theta_0 - \bar{\theta}\|_2^2 + \|q_0\|_2^2 < \delta^2, \tag{2.1}$$

the solution of (1.14), (1.17), (1.18) decays exponentially as $t \to +\infty$.

In order to carry out the proof, we consider another problem which agrees with (1.14), (1.17), (1.18) when (θ, q) are close enough to the equilibrium state $(\bar{\theta}, 0)$. For this purpose, we introduce the functions A, B, C, D satisfying the following hypotheses

h1) $A, B \in C_b^2(\mathbb{R}^2)$ and $C, D, E \in C_b^2(\mathbb{R}^2)$ such that

$$\begin{aligned} A(\xi) &= \frac{\kappa(\xi)}{\tau(\xi)}, \qquad B(\xi) = \frac{1}{\tau(\xi)}, \qquad C(\xi,\zeta) = \frac{1}{e_0'(\xi) + a'(\xi)\zeta^2}\\ D(\xi,\zeta) &= \frac{2a(\xi)\kappa(\xi)}{(e_0'(\xi) + a'(\xi)\zeta^2)\tau(\xi)}, \qquad E(\xi,\zeta) = \frac{2a(\xi)\kappa(\xi)\zeta}{(e_0'(\xi) + a'(\xi)\zeta^2)\tau(\xi)} \end{aligned}$$

$\forall(\xi,\zeta) \in \mathcal{V}$.

h2) $A(\xi) \geq \underline{A} > 0, \quad B(\xi) \geq \underline{B} > 0, C(\xi,\eta) \geq \underline{C} > 0.$.

Here C_b^2 denotes the space of continuous and bounded functions, as well as, their first and second order derivatives. We note that functions with these properties can

be constructed by virtue of (1.13). Therefore, instead of (1.14), (1.17), (1.18), we consider the following problem

$$q_t = -A(\theta)\theta_x - B(\theta)q \tag{2.2}$$
$$\theta_t = -C(\theta)q_x + D(\theta,q)q^2 + E(\theta,q)\theta_x, \quad x \in I \quad t > 0. \tag{2.3}$$
$$\theta(x,0) = \theta_0(x),\ q(x,0) = q_0(x), \quad x \in I \tag{2.4}$$
$$\theta(0,t) - \bar{\theta} = \theta(1,t) - \bar{\theta} = 0, \quad t \geq 0 \tag{2.5}$$

Remark. If (θ, q) is a solution of (2.2) – (2.5), which remains in $\mathcal{V}$ then by virtue of (h2), it is also a solution to (1.14), (1.17), (1.18).

We set

$$\mathcal{E}(t) \ : \ = \int_{-\infty}^{\infty} [(\theta - \bar{\theta})^2 + \theta_t^2 + \theta_x^2 + \theta_{xt}^2 + \theta_{xx}^2 + \theta_{tt}^2 + q^2 + q_t^2 + q_x^2 + q_{xt}^2 + q_{xx}^2 + q_{tt}^2](x,t)dx, \tag{2.6}$$

$$\Lambda(t) \ : \ = \int_{-\infty}^{\infty} A(\theta)[(\theta - \bar{\theta})^2 + \theta_t^2 + \theta_{tt}^2](x,t)dx + \int_{-\infty}^{\infty} C(\theta)[q^2 + q_t^2 + q_{tt}^2)(x,t)dx, \tag{2.7}$$

and

$$\alpha(t) := \max_x \Big[(|\theta - \bar{\theta}\,| + |\theta_x| + |\theta_t| + |q| + |q_t| + |q_x|)(x,t)\Big] \tag{2.8}$$

Proof.

We multiply (2.2) by $C(\theta)q$ and (2.3) by $A(\theta)(\theta - \bar{\theta})$ integrate over I, use integration by parts, and add equalities, to get

$$\frac{d}{dt}\{\frac{1}{2}\int_0^1 [A(\theta - \bar{\theta})^2 + Cq^2]dx\} \leq -\int_0^1 CBq^2 + \Gamma\alpha(t)\mathcal{E}(t) \tag{2.9}$$

where Γ is a generic positive constant independent of θ, q, and t. We then differentiate (2.2), (2.3) with respect to t ; hence we have

$$q_{tt} = -A\theta_{xt} - A_t\theta_x - Bq_t - B_tq \tag{2.10}$$
$$\theta_{tt} = -Cq_{xt} - C_tq_x + D_tq^2 + 2Dqq_t + E_t\theta_x + E\theta_{xt}. \tag{2.11}$$

We multiply (2.10) by Cq_t and (2.11) by $A\theta_t$ integrate over I, use integration by parts, and add equalities, to obtain

$$\frac{d}{dt}\{\frac{1}{2}\int_0^1 [A\theta_t^2 + Cq_t^2]dx\} \leq -\int_0^1 CBq_t^2 + \Gamma[\alpha(t) + \alpha^2(t)]\mathcal{E}(t). \tag{2.12}$$

To establish bounds on terms involving θ_{tt} and q_{tt}, we introduce a difference operator Δ_h as follows : for $h > 0$, we set

$$\Delta_h W(x,t) = W(x,t+h) - W(x,t), \quad x \in I, \quad t \geq 0. \tag{2.13}$$

We apply the above operator to equations (2.10), (2.11), multiply the resulting equalities by $C(\theta)\Delta_h q_t$ and $A\Delta_h\theta_t$ respectively, integrate over I, and add the inequalities. After a number of integrations, using integration by parts, we divide by h^2 and let h go to zero. Thus we get

$$\begin{aligned} \frac{d}{dt}\{\frac{1}{2}\int_0^1 [A\theta_{tt}^2 + Cq_{tt}^2]dx\} &\leq -\int_0^1 CBq_{tt}^2 \\ &+\Gamma[\alpha(t)+\alpha^2(t)+\alpha^3(t)]\mathcal{E}(t). \end{aligned} \tag{2.14}$$

Therefore, by combining (2.9), (2.12), and (2.14) we obtain

$$\Lambda' \leq -2\int_0^1 CB[q^2+q_t^2+q_{tt}^2]dx + \Gamma[\alpha(t)+\alpha^2(t)+\alpha^4(t)]\mathcal{E}(t). \tag{2.15}$$

Next we show that, for $(\theta, q) \in \mathcal{V}$, Λ is equivalent to $\mathcal{E}$. For this purpose we use equations (2.2), (2.3), (2.10), (2.11), and the hypotheses (h1), (h2) to arrive at

$$\begin{aligned} \int_0^1 [\theta_x^2 + q_x^2 + \theta_{xt}^2 + q_{xt}^2]dx &\leq c_1\int_0^1 [q^2+q_t^2+q_{tt}^2+dx+\theta_t^2+\theta_{tt}^2]dx \\ &+c_1[1+\alpha^2(t)+\alpha^4(t)]\Lambda(t). \end{aligned} \tag{2.16}$$

We then differentiate (2.2), (2.3) with respect to x and use the resulting equations to get

$$\int_0^1 [\theta_{xx}^2 + q_{xx}^2]dx \leq c_1[1+\alpha^2(t)+\alpha^4(t)]\Lambda(t) \tag{2.17}$$

Therefore, a combination of (2.16) and (2.17) yields

$$c_1\Lambda(t) \leq \mathcal{E}(t) \leq c_2\{1+\alpha^2(t)+\alpha^4(t)\}\Lambda(t) \tag{2.18}$$

where c_1, c_2 are constants independent of t; hence (2.15) takes the form

$$\Lambda'(t) \leq -2a\int_0^1 \left(q^2+q_t^2+q_{tt}^2\right)dx + \Gamma\alpha(t)\{1+\alpha^7(t)\}\Lambda(t). \tag{2.19}$$

For further estimate, we multiply (2.3) by θ_t and integrate over I; so we have

$$\begin{aligned} \int_0^1 \theta_t^2 dx &\leq -\int_0^1 Cq_x\theta_t dx + \Gamma\alpha(t)\{1+\alpha^7(t)\}\Lambda(t) \\ &\leq \int_0^1 Cq\theta_{xt}dx + \Gamma\alpha(t)\{1+\alpha^7(t)\}\Lambda(t) \\ &\leq \frac{d}{dt}\int_0^1 Cq\theta_x dx - \int_0^1 Cq_t\theta_x dx + \Gamma\alpha(t)\{1+\alpha^7(t)\}\Lambda(t) \end{aligned} \tag{2.20}$$

which implies, by virtue of (2.2), that

$$\int_0^1 \theta_t^2 dx - \frac{d}{dt}\int_0^1 Cq\theta_x dx \leq c\int_{-\infty}^{\infty}\left(q^2+q_t^2\right)dx + \Gamma\alpha(t)\{1+\alpha^7(t)\}\Lambda(t). \tag{2.21}$$

A similar treatment to (2.11) leads to

$$\int_0^1 \theta_{tt}^2 dx - \frac{d}{dt}\int_0^1 Cq_t\theta_{xt} dx \leq c\int_{-\infty}^{\infty}\left(q_t^2+q_{tt}^2\right)dx + \Gamma\alpha(t)\{1+\alpha^7(t)\}\Lambda(t). \tag{2.22}$$

Also, Poincaré ' s inequality and equation (2.2) give

$$\int_0^1 (\theta - \bar{\theta})^2 dx \leq c \int_{-\infty}^{\infty} \left(q^2 + q_t^2\right) dx \tag{2.23}$$

Thus (2.21), (2.22) and (2.23) yield

$$\int_0^1 [(\theta - \bar{\theta})^2 + \theta_t^2 + \theta_{tt}^2] dx - G^{'}(t) \leq c \int_{-\infty}^{\infty} \left(q^2 + q_t^2 + q_{tt}^2\right) dx$$
$$+ \Gamma \alpha(t)\{1 + \alpha^7(t)\}\Lambda(t), \tag{2.24}$$

where

$$G(t) = \int_0^1 C[q\theta_x + q_t \theta_{xt}] dx$$

We also define $F = \Lambda - \varepsilon G$, for ε so small that, by virtue of the definitions of Λ, G, (h1), and (h2), we obtain

$$c_3 \Lambda(t) \leq F(t) \leq c_4 \Lambda(t)$$

where c_3, c_4 are constants independent of t. Therefore a combination of (2.15) and (2.24), using (h2) and choosing ε as small as necessary, leads to

$$F^{'}(t) \leq -bF(t) + \Gamma \alpha(t)\{1 + \alpha^7(t)\}F(t), \tag{2.25}$$

for some constant $b > 0$. We also note that, by the standard Sobolev embedding inequalities we have $\alpha(t) \leq \sqrt{2\mathcal{E}(t)}$. So by choosing δ in (2.1) so small that $(\theta_0, q_0) \in \mathcal{V}$ and $\Gamma\alpha(0)\{1 + \alpha^7(0)\} < b/2$, the continuity yields $(\theta, q) \in \mathcal{V}$ and $\Gamma\alpha(t)\{1 + \alpha^7(t)\} < b/2$, $\forall t \in [0, t_0)$; consequently relation (2.25) yields

$$F'(t) < -\frac{b}{2} F(t), \qquad \forall t \in [0, t_0) \tag{2.26}$$

Direct integration of (2.26) over $(0, t)$ then yields

$$F(t) \leq F(0) e^{-bt/2}, \qquad \forall t \in [0, t_0); \tag{2.27}$$

hence $F(t) \leq F(0)$ and we can extend (2.27) beyond t_0. By repeating the same procedure, taking δ even smaller if necessary, and using the continuity of F, (2.27) is established for all $t \geq 0$. This completes the proof.

Acknowledgment This work has been funded by King Fahd University of Petroleum & Minerals under Project # MS/SOUND/238.

References

1. C. Cattaneo, Sulla conduzione del calore, *Atti Sem. Math. Fis Univ. Modena* 3, 83-101 (1948).

2. B. D. Coleman and V. Mizel, Thermodynamic and departure from Fourier's law of heat conduction, *Arch. Rational Mech. Anal.* 13, 245 - 261 (1963).

3. B. D. Coleman, W. J. Hrusa, and D. R. Owen, Stability of equilibrium for a nonlinear hyperbolic system describing heat propagation by second sound in solids, *Arch. Rational Mech. Anal.* 94, 267–289 (1986).

4. B. D. Coleman, M. Fabrizio, and D. R. Owen, On the thermodynamics of second sound in dielectric crystals, *Arch. Rational Mech. Analysis* 80, 135–158 (1982).

5. S. A. Messaoudi, Formation of singularities in heat propagation guided by second sound, *J. Diff. Eqns.* 130, 92–99 (1996).

6. S. A. Messaoudi, On the existence and nonexistence of solutions of a nonlinear hyperbolic system describing heat propagation by second sound, *Applicable Analysis* 73, 485-496 (1999).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,257-263,
2004,

Necessary conditions for the matrix equation $AXB + CXD = E$ to be consistent

Yongge Tian
Department of Mathematics and Statistics, Queen's University
Kingston, Ontario, Canada K7L 3N6 E-mail: ytian@mast.queensu.ca

Abstract. Matrix equivalence can be used to characterize consistency of linear matrix equations. The well-known Roth's equivalence theorem states that the matrix equation $AX + YB = C$ is solvable for X and Y if and only if $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ and $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ are equivalent, i.e., $\operatorname{rank}\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \operatorname{rank}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. In this note, we partially extend the work to the matrix equation $AXB+CXD = E$ and show a group of rank equalities which are necessary for $AXB + CXD = E$ to be consistent.

AMS subject classifications: 15A24.
Key words: matrix equivalence, rank equality, matrix equation, consistency.

In matrix theory it is well known that the matrix equation $AX - YB = C$ is solvable for X and Y if and only if

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \text{ is equivalent to } \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}, \tag{1}$$

or equivalently

$$\operatorname{rank}\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \operatorname{rank}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}, \tag{2}$$

see [14]. This result is often called Roth's equivalence theorem, and was revisited by lots of authors, see, e.g., [3, 5, 6, 7, 11, 16]). Roth also showed in [14] that the Sylvester equation $AX - XB = C$ is solvable for X if and only if

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \text{ is similar to } \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}. \tag{3}$$

This result is often called Roth's removal theorem.

Roth's work was extented to various general settings. For example, it was shown by Özgüler [13] that there exist X and Y such that $AXB + CYD = E$ if and only if the four independent rank equalities

$$r[\,A,\ C,\ E\,] = r[\,A,\ C],\qquad r\begin{bmatrix} B \\ D \\ E \end{bmatrix} = r\begin{bmatrix} B \\ D \end{bmatrix}, \tag{4}$$

$$r\begin{bmatrix} A & E \\ 0 & D \end{bmatrix} = r(A) + r(D), \qquad r\begin{bmatrix} C & E \\ 0 & B \end{bmatrix} = r(C) + r(B) \tag{5}$$

hold, where $r(\cdot)$ denotes the rank of a matrix. This work is extented to the linear matrix equation $A_1X_1B_1 + A_2X_2B_2 + A_3X_3B_3 = C$ by the author [15]. On the other hand, the consistency and solution of the matrix equation

$$AXB + CXD = E \tag{6}$$

was also examined through various methods, see, e.g., [1, 2, 4, 8, 10, 12].

A powerful method in solving linear matrix equations is the well-known vec operation and Kronecker product of matrices. Any linear matrix equation can be converted through the method to a trivial form $Mx = b$ (see [9]). However, it is difficult in general to remove the vec operation and Kronecker product from the consistency condition and solution of $Mx = b$. Hence it is unknown at present how to characterize the consistency of the equation (6) using equivalence and similarity of matrices.

In this note, we show some partial results associated with the consistency of the equation (6).

Theorem. *Let A, C be $m \times p$ matrices, B, D be $q \times n$ matrices, and E be an $m \times n$ matrix over an arbitrary field $\mathbb{F}$. If there is an matrix X satisfying the equation* (6), *then A, B, C, D and E satisfy the following rank equalities*

$$r[\,A,\ C,\ E\,] = r[\,A,\ C,\ 0\,], \qquad r\begin{bmatrix} B \\ D \\ E \end{bmatrix} = r\begin{bmatrix} B \\ D \\ 0 \end{bmatrix}, \tag{7}$$

$$r\begin{bmatrix} A & E \\ 0 & D \end{bmatrix} = r\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}, \quad r\begin{bmatrix} C & E \\ 0 & B \end{bmatrix} = r\begin{bmatrix} C & 0 \\ 0 & B \end{bmatrix}, \tag{8}$$

$$r\begin{bmatrix} M_{1k} & E_k \\ 0 & N_{1k} \end{bmatrix} = r\begin{bmatrix} M_{1k} & 0 \\ 0 & N_{1k} \end{bmatrix}, \quad r\begin{bmatrix} M_{2k} & E_k \\ 0 & N_{2k} \end{bmatrix} = r\begin{bmatrix} M_{2k} & 0 \\ 0 & N_{2k} \end{bmatrix}, \tag{9}$$

$$r\begin{bmatrix} M_{3k} & E_k \\ 0 & N_{3k} \end{bmatrix} = r\begin{bmatrix} M_{3k} & 0 \\ 0 & N_{3k} \end{bmatrix}, \quad r\begin{bmatrix} M_{4k} & E_k \\ 0 & N_{4k} \end{bmatrix} = r\begin{bmatrix} M_{4k} & 0 \\ 0 & N_{4k} \end{bmatrix}, \tag{10}$$

for $k = 2,\ 3,\ \ldots,$ where

$$E_k = \begin{bmatrix} E & & & \\ & -E & & \\ & & \ddots & \\ & & & (-1)^{k+1}E \end{bmatrix}_{k \times k}, \quad M_{1k} = \begin{bmatrix} A & C & & & \\ & A & C & & \\ & & \ddots & \ddots & \\ & & & A & C \end{bmatrix}_{k \times (k+1)},$$

$$N_{2k} = \begin{bmatrix} D & B & & & \\ & D & B & & \\ & & \ddots & \ddots & \\ & & & D & B \end{bmatrix}_{(k-1) \times k}, \quad M_{2k} = \begin{bmatrix} A & & & \\ C & A & & \\ & C & \ddots & \\ & & \ddots & A \\ & & & C \end{bmatrix}_{k \times (k-1)},$$

$$N_{2k} = \begin{bmatrix} D & & & \\ B & D & & \\ & B & \ddots & \\ & & \ddots & D \\ & & & B \end{bmatrix}_{(k+1)\times k}, \quad M_{3k} = \begin{bmatrix} A & C & & \\ & A & \ddots & \\ & & \ddots & C \\ & & & A \end{bmatrix}_{k\times k},$$

$$N_{3k} = \begin{bmatrix} D & B & & \\ & D & \ddots & \\ & & \ddots & B \\ & & & D \end{bmatrix}_{k\times k}, \quad M_{4k} = \begin{bmatrix} C & A & & \\ & C & \ddots & \\ & & \ddots & A \\ & & & C \end{bmatrix}_{k\times k},$$

$$N_{4k} = \begin{bmatrix} B & D & & \\ & B & \ddots & \\ & & \ddots & D \\ & & & B \end{bmatrix}_{k\times k}.$$

Proof. The four rank equalities in (7) and (8) come directly from (4) and (5), because the consistency of the equation (6) implies the consistency of $AXB + CYD = E$. We next show that the rank equalities in (9) and (10) hold for any k. Suppose the equation (6) has a solution X_0, that is, $AX_0B + CX_0D = E$. In this case, construct two nonsingular block matrices as follows

$$P_{1k} = \begin{bmatrix} I_m & & & & & 0 & & & \\ & I_m & & & & AX_0 & & & \\ & & I_m & & & & -AX_0 & & \\ & & & \ddots & & & & \ddots & \\ & & & & I_m & & & & (-1)^k AX_0 \\ & & & & & I_q & & & \\ & & & & & & I_q & & \\ & & & & & & & \ddots & \\ & & & & & & & & I_q \end{bmatrix}_{(2k-1)\times(2k-1)},$$

$$Q_{1k} = \begin{bmatrix} I_p & & & & & -X_0B & & & \\ & I_p & & & & -X_0D & & & \\ & & I_p & & & & X_0D & & \\ & & & \ddots & & & & \ddots & \\ & & & & I_p & & & & (-1)^k X_0D \\ & & & & & I_n & & & \\ & & & & & & I_n & & \\ & & & & & & & \ddots & \\ & & & & & & & & I_n \end{bmatrix}_{(2k+1)\times(2k+1)},$$

and calculate to yield the equality

$$P_{1k}\begin{bmatrix} M_{1k} & E_k \\ 0 & N_{1k} \end{bmatrix} Q_{1k} = \begin{bmatrix} M_{1k} & 0 \\ 0 & N_{1k} \end{bmatrix},$$

which is equivalent to the first rank equality in (9). Let

$$P_{3k} = \begin{bmatrix} I_m & & & & & 0 & & & & \\ & I_m & & & & AX_0 & & & & \\ & & I_m & & & & -AX_0 & & & \\ & & & \ddots & & & & \ddots & & \\ & & & & I_m & & & & (-1)^k AX_0 & (-1)^k CX_0 \\ & & & & & I_q & & & & \\ & & & & & & I_q & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & I_q & \\ & & & & & & & & & I_q \end{bmatrix}_{2k\times 2k},$$

$$Q_{3k} = \begin{bmatrix} I_p & & & & & -X_0B & & & & \\ & I_p & & & & -X_0D & & & & \\ & & I_p & & & & X_0D & & & \\ & & & \ddots & & & & \ddots & & \\ & & & & I_p & & & & (-1)^{k-1}X_0D & 0 \\ & & & & & I_n & & & & \\ & & & & & & I_n & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & I_n & \\ & & & & & & & & & I_n \end{bmatrix}_{2k\times 2k},$$

and calculate to yield the equality

$$P_{3k}\begin{bmatrix} M_{3k} & E_k \\ 0 & N_{3k} \end{bmatrix} Q_{3k} = \begin{bmatrix} M_{3k} & 0 \\ 0 & N_{3k} \end{bmatrix},$$

which is the first rank equality in (10). The other two rank equalities in (9) and (10) can be shown by a similar approach. Their proofs are omitted here. □

When $k = 2$, the four rank equalities in (9) and (10) become

$$r\begin{bmatrix} A & C & 0 & E & 0 \\ 0 & A & C & 0 & -E \\ 0 & 0 & 0 & D & B \end{bmatrix} = r\begin{bmatrix} A & C & 0 & 0 & 0 \\ 0 & A & C & 0 & 0 \\ 0 & 0 & 0 & D & B \end{bmatrix}, \tag{11}$$

$$r\begin{bmatrix} A & E & 0 \\ C & 0 & -E \\ 0 & D & 0 \\ 0 & B & D \\ 0 & 0 & B \end{bmatrix} = r\begin{bmatrix} A & 0 & 0 \\ C & 0 & 0 \\ 0 & D & 0 \\ 0 & B & D \\ 0 & 0 & B \end{bmatrix}, \tag{12}$$

$$r\begin{bmatrix} A & C & E & 0 \\ 0 & A & 0 & -E \\ 0 & 0 & D & B \\ 0 & 0 & 0 & D \end{bmatrix} = r\begin{bmatrix} A & C & 0 & 0 \\ 0 & A & 0 & 0 \\ 0 & 0 & D & B \\ 0 & 0 & 0 & D \end{bmatrix}, \tag{13}$$

$$r\begin{bmatrix} C & A & E & 0 \\ 0 & C & 0 & -E \\ 0 & 0 & B & D \\ 0 & 0 & 0 & B \end{bmatrix} = r\begin{bmatrix} C & A & 0 & 0 \\ 0 & C & 0 & 0 \\ 0 & 0 & B & D \\ 0 & 0 & 0 & B \end{bmatrix}. \tag{14}$$

When $k = 3$, the four rank equalities in (9) and (10) become

$$r\begin{bmatrix} A & C & 0 & 0 & E & 0 & 0 \\ 0 & A & C & 0 & 0 & -E & 0 \\ 0 & 0 & A & C & 0 & 0 & E \\ 0 & 0 & 0 & 0 & D & B & 0 \\ 0 & 0 & 0 & 0 & 0 & D & B \end{bmatrix} = r\begin{bmatrix} A & C & 0 & 0 & 0 & 0 & 0 \\ 0 & A & C & 0 & 0 & 0 & 0 \\ 0 & 0 & A & C & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & D & B & 0 \\ 0 & 0 & 0 & 0 & 0 & D & B \end{bmatrix}, \tag{15}$$

$$r\begin{bmatrix} A & 0 & E & 0 & 0 \\ C & A & 0 & -E & 0 \\ 0 & C & 0 & 0 & E \\ 0 & 0 & D & 0 & 0 \\ 0 & 0 & B & D & 0 \\ 0 & 0 & 0 & B & D \\ 0 & 0 & 0 & 0 & B \end{bmatrix} = r\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ C & A & 0 & 0 & 0 \\ 0 & C & 0 & 0 & 0 \\ 0 & 0 & D & 0 & 0 \\ 0 & 0 & B & D & 0 \\ 0 & 0 & 0 & B & D \\ 0 & 0 & 0 & 0 & B \end{bmatrix}, \tag{16}$$

$$r\begin{bmatrix} A & C & 0 & E & 0 & 0 \\ 0 & A & C & 0 & -E & 0 \\ 0 & 0 & A & 0 & 0 & E \\ 0 & 0 & 0 & D & B & 0 \\ 0 & 0 & 0 & 0 & D & B \\ 0 & 0 & 0 & 0 & 0 & D \end{bmatrix} = r\begin{bmatrix} A & C & 0 & 0 & 0 & 0 \\ 0 & A & C & 0 & 0 & 0 \\ 0 & 0 & A & 0 & 0 & 0 \\ 0 & 0 & 0 & D & B & 0 \\ 0 & 0 & 0 & 0 & D & B \\ 0 & 0 & 0 & 0 & 0 & D \end{bmatrix}, \tag{17}$$

$$r\begin{bmatrix} C & A & 0 & E & 0 & 0 \\ 0 & C & A & 0 & -E & 0 \\ 0 & 0 & C & 0 & 0 & E \\ 0 & 0 & 0 & B & D & 0 \\ 0 & 0 & 0 & 0 & B & D \\ 0 & 0 & 0 & 0 & 0 & B \end{bmatrix} = r\begin{bmatrix} C & A & 0 & 0 & 0 & 0 \\ 0 & C & A & 0 & 0 & 0 \\ 0 & 0 & C & 0 & 0 & 0 \\ 0 & 0 & 0 & B & D & 0 \\ 0 & 0 & 0 & 0 & B & D \\ 0 & 0 & 0 & 0 & 0 & B \end{bmatrix}. \tag{18}$$

Since there are infinitely-many rank equalities in (9) and (10), it is needed to consider the independence of these rank equalities. It can be seen from the structure of (7)–(10) that for any given $k \geq 2$, the six rank equalities in them are independent in general, that is to say, any one or some of six rank equalities do not imply other rank equalities. Moreover, we guess that for all $k \leq \min\{p,\ q\}$, the equivalences in (7)–(10) are all independent. But for $k > \min\{p,\ q\}$, nothing can be said about the independence of the rank equalities in (9) and (10).

The discovery of the four types of rank equalities in (9) and (10) is somewhat of surprise for the equation (6). As a matter of fact, if any one of these equivalences is not satisfied, then (6) is not solvable for X. Thus when examining the consistency of (6), one should sufficiently consider the rank equalities in (9) and (10), although they are not sufficient conditions for (6) to be consistent. Following this remarks, it is worth considering the following two problems:

(I) What conditions together with (7)–(10) are necessary and sufficient for (6) to be consistent?

(II) Under what conditions, the rank equalities in (7)–(10) are sufficient for (6) to be consistent?

For Problem (I), we can say nothing at present. Although it is best known that the equation (6) is solvable if and only if the conventional system of linear equations $[\,(B^T \otimes A) + (D^T \otimes C)\,]\overrightarrow{X} = \overrightarrow{E}$ is solvable, one can hardly establish any essential relationship between (7)–(10) and the solvability of the system of equations. For Problem (II), we have the following conjectures.

Conjecture 1. Under the condition

$$\mathscr{R}(A) \cap \mathscr{R}(C) = \{0\}, \quad \text{or} \quad \mathscr{R}(B^T) \cap \mathscr{R}(D^T) = \{0\}, \tag{19}$$

where $\mathscr{R}(\cdot)$ denotes the range (column space) of a matrix, there is an X such that $AXB + CXD = E$ if and only if the eight rank equalities in (7), (8), (11)–(14) hold.

Conjecture 2. Under the condition

$$r\begin{bmatrix} A & C & 0 \\ 0 & A & C \end{bmatrix} = r\begin{bmatrix} A & C \\ 0 & A \end{bmatrix} + r\begin{bmatrix} C & 0 \\ A & C \end{bmatrix} - r\begin{bmatrix} C \\ A \end{bmatrix}, \tag{20}$$

or

$$r\begin{bmatrix} B & 0 \\ D & B \\ 0 & D \end{bmatrix} = r\begin{bmatrix} B & 0 \\ D & B \end{bmatrix} + r\begin{bmatrix} D & B \\ 0 & D \end{bmatrix} - r[\,D,\ B\,], \tag{21}$$

there is an X such that $AXB + CXD = E$ if and only if the twelve rank equalities in (7), (8) and (11)–(18) hold.

When both B and C are identity matrices in the equation (6), Conjecture 2 reduces to the following form.

Conjecture 3. Under the condition $A^2 = 0$ or $D^2 = 0$, the matrix equation $AX - XD = E$ is solvable if and only if

$$r\begin{bmatrix} A & E \\ 0 & D \end{bmatrix} = r\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}, \ r\begin{bmatrix} A & E \\ 0 & D \end{bmatrix}^2 = r\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}^2, \ r\begin{bmatrix} A & E \\ 0 & D \end{bmatrix}^3 = r\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}^3, \tag{22}$$

or equivalently, the two statements in (3) and (22) are equivalent.

References

[1] K.E. Chu, The solutions of the matrix equations $AXB - CXD = E$ and $(YA - DZ, YC - BZ) = (E,\ F)$, Linear Algebra Appl. 93, 93–105(1987).
[2] M.A. Epton, Methods for the solution of $AXB - CXD = E$ and its application in the numerical solution of implicit ordinary differential equations, BIT 20, 341–345(1980).
[3] H. Flanders and H.K. Wimmer, On the matrix equations $AX - XB = C$ and $AX - YB = C$, SIAM J. Appl. Anal. Math. 32, 707–710(1977).
[4] J.D. Gardiner, A.J. Laub, J.J. Amato and C.B. Moler, Solution of the Sylvester matrix equation $AXB^T + CXD^T = E$, ACM Trans. Math. Software 18, 223–231(1992).
[5] W.H. Gustafson, Roth's theorems over commutative rings, Linear Algebra Appl. 23, 245–251(1979).
[6] W.H. Gustafson and J. M. Zelmanowitz, On matrix equivalence and matrix equations, Linear Algebra Appl. 27, 219–224(1979).
[7] R.E. Hartwig, Roth's equivalence problems in unit regular rings, Proc. Amer. Math. Soc. 59, 39–44(1976).
[8] V. Hernández and M. Gassó, Explicit solution of the matrix equation $AXB - CXD = E$, Linear Algebra Appl. 121, 333–344(1989).
[9] R. A. Horn and C. R. Johnson, Topics in matrix analysis, Cambridge University Press, Cambridge, 1991.
[10] P. Lancaster, Explicit solutions of linear matrix equations, SIAM Review 12, 544–566(1970).
[11] J. Li, On a proof of Roth's theorem, J. Math. Res. Exposition 4, 75–76(1984).
[12] S.K. Mitra, The matrix equation $AXB + CXD = E$, SIAM J. Appl. Anal. Math. 32, 823–825(1977).
[13] A.B. Özgüler, The matrix equation $AXB + CYD = E$ over a principal ideal domain, SIAM J. Matrix. Anal. Appl. 12, 581–591(1991).
[14] W.E. Roth, The equations $AX - YB = C$ and $AX - XB = C$ in matrices, Proc. Amer. Math. Soc. 3, 392–396(1952).
[15] Y. Tian, Solvability of two linear matrix equations, Linear and Multilinear Algebra 48, 123–147(2000).
[16] A.J.B. Ward, A straightforward proof of Roth's lemma in matrix equations, Internat. J. Math. Ed. Sci. Tech. 30, 33–38(1999).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,265-279,
2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Parameter Estimation in a Nonlinear Phytoplankton Aggregation Model

Robert R. Ferdinand
Department of Mathematics
East Central University
Ada, OK 74820-6899

February 14, 2006

Abstract

In this paper, a model describing phytoplankton population dynamics in an environment is presented. An implicit finite difference scheme is used to approximate solution of this model. Then an inverse method procedure, involving the minimization of a least squares cost functional, is used to estimate certain model parameters using computationally generated data. Further, convergence results for these parameter estimates are established and finally, numerical examples are presented which confirm the theoretical results proved.

Mathematics Subject Classification Numbers: 35L60, 35L65, 65K10, 65N06, 65N12, 65N21
Key Words: Finite Difference, Hyperbolic PDE, Inverse Problem, Parameter Estimation

1 Introduction

The nonlinear, nonlocal, hyperbolic partial differential equation model (1.1) shown below, describes dynamics of phytoplankton population with aggregation in an ocean environment.

$$\begin{cases} u_t + (g(x,P(t))u)_x = F(u) - w(x)u & (t,x) \in (0,T] \times (0,x_{\max}] \\ g(0,P(t))u(0,t) = \displaystyle\int_0^{x_{\max}} \beta(x,P(t))u(t,x)dx & t \in (0,T] \\ u(0,x) = u^0(x) & x \in [0,x_{\max}] \end{cases} \tag{1.1}$$

where

$$F(u) = \frac{1}{2}\int_0^x \eta(x-y,y)u(t,x-y)u(t,y)dy - u(t,x)\int_0^{x_{\max}} \eta(x,y)u(t,y)dy$$

is the nonlinear aggregation term and

$$P(t) = \int_0^{x_{\max}} u(t,x)dx$$

represents total phytoplankton population at time t. $u(t,x)$ denotes population density of phytoplankton having size x at time t, while parameters g and β denote growth and reproduction rates of an individual, respectively. It may be noted that both g and β are functions of size x and total population P. $w(x)$ is the sinking/mortality rate of phytoplankton having size x, while $\eta(x,y)$ is the rate at which an individual of size

x coalesces with an individual of size y, when they come into contact upon collision. Finally, u^0 is the initial population distribution.

It may be worthwhile to mention that the above model contains an all important nonlinear aggregation/reaction term $F(u)$ which represents the coalescing of particles upon collision in an ocean. This causes the parameter estimation problem associated with this model to be different from any of those discussed in [1, 2, 6, 7, 8, 14].

This particular model (1.1) was first presented in [4]. In that paper, an implicit finite difference scheme was used to approximate model solution and convergence of the numerical approximants to a unique, bounded variation, weak solution was shown.

In (1.1), parameters such as the growth and reproduction functions represent physiologically significant processes in the life cycle of phytoplankton. Hence, a need arises for one to be able to numerically estimate parameters such as these.

The procedure employed to identify function parameters involves comparing observed phytoplankton densities with solution output of the model and minimizing a least-squares cost functional in the process. A minimizer to this cost functional is computed numerically and its convergence to a minimizer of the original least-squares cost functional is shown theoretically. This method has been implemented successfully in several other phytoplankton population models (see [1, 2, 6, 7, 8, 14]).

Hence the following parameter identification problem presents itself:

- Given observations Π_r, corresponding to total population of phytoplankton in the interval $[0, x_{\max}]$ at time t_r, $r = 0, \cdots, R$, find a parameter $q = \left(g,\ \beta,\ w,\ \eta,\ u^0\right) \in Q$ which minimizes the following least squares cost functional or

$$\min_{q \in Q} J(q) = \sum_{r=0}^{R} \left| \int_0^{x_{\max}} u(t_r, x, q)dx - \Pi_r \right|^2, \tag{1.2}$$

  where $u(t, x, q)$ represents the parameter dependent solution of model equation (1.1) and Q represents an infinite dimensional parameter space.

This paper is organized as follows. In section 2, an inverse problem approximation method which numerically estimates parameter q is described, while section 3 deals with proving convergence results of computed minimizers to a minimizer of the original least squares problem. In section 4, numerical results are presented and future research issues are addressed in section 5.

2 Approximating Parameters Numerically using Inverse Method

The following regularity conditions are imposed on the infinite dimensional parameter space Q.

(B_Q) We define the space

$$D \quad = \quad C^{1,0}(\Omega) \times C(\Omega) \times C[0, x_{\max}] \times C([0, x_{\max}] \times [0, x_{\max}]) \times L^1(0, x_{\max}).$$

Here $\Omega = [0, x_{\max}] \times [0, \infty)$. The admissible parameter space Q is thus a compact subset of D which satisfies the following conditions uniformly in q for all $q \in Q$.

(B_g) Growth function $g(x, P)$ is twice continuously differentiable with respect to x and continuously differentiable with respect to P. Also, $g > 0$ for $x \in [0, x_{\max})$ and $g(x_{\max}, P) = 0$. Finally,

$$|g| + |g_x| + |g_P| + |g_{xx}| + |g_{xP}| \leq A_1,$$

where A_1 is a fixed constant.

(B_β) $\beta(x, P)$ is non-negative and continuously differentiable with respect to x and P. Further,

$$|\beta| + |\beta_x| + |\beta_P| \leq A_2,$$

A_2 being a fixed constant.

(B_w) $w(x)$ is a non-negative continuously differentiable function and $|w(x)| \leq A_3$, where A_3 is a fixed constant.

(B_η) $\eta(x, y)$, the aggregation kernel, is non-negative and continuously differentiable with $\|\eta\|_{C^1} \leq A_4$. Further,

$$\begin{cases} \eta(x,y) \geq 0 & ; \quad x + y \leq x_{\max} \\ \eta(x,y) = 0 & ; \quad x + y > x_{\max}. \end{cases}$$

(B_{u^0}) Initial condition $u^0(x) \in BV(0, x_{\max}) \cap L^\infty(0, x_{\max})$ and $u^0(x) \geq 0$.

In this paper, techniques similar to those used in [3, 13] will be used to establish convergence of computed minimizers for the inverse problem (1.2) and (1.1).

We begin by defining a parameter dependent weak solution $u(t, x, q)$ to the partial differential equation model (1.1) as a bounded measurable function satisfying

$$\begin{aligned} &\int_0^{x_{\max}} u(t,x,q)\phi(t,x)dx - \int_0^{x_{\max}} u^0(x)\phi(0,x)dx \\ &= \int_0^t \int_0^{x_{\max}} u\left(\phi_s + g\phi_x - w\phi\right) dxds \\ &+ \int_0^t \phi(s,0) \left(\int_0^{x_{\max}} \beta(x, P(s,q))u(s,x,q)dx \right) ds \\ &+ \int_0^t \int_0^{x_{\max}} F\left(u\left(s,x,q\right)\right)\phi dxds, \end{aligned} \tag{2.1}$$

for $t \in [0, T]$ and each function $\phi \in C^1((0, T) \times (0, x_{\max}))$.

The first step taken while solving the least squares problem involves approximating the solution of equation (1.1) numerically, which is carried out as follows: Let

$$\Delta x = \frac{x_{\max}}{N} \quad \text{and} \quad \Delta t = \frac{T}{\Theta}$$

represent the space and time mesh sizes, respectively. The mesh points are given as $x_j = j\Delta x$, $j = 0, \cdots, N$ and $t_k = k\Delta t$, $k = 0, \cdots, \Theta$. Also, the finite difference approximations of $u(t_k, x_j, q)$ and $P(t_k, q)$ are denoted by $u_j^k(q)$ and $P^k(q)$, respectively. Further,

$$g_j^k = g\left(x_j, P^k(q)\right), \ \beta_j^k = \beta\left(x_j, P^k(q)\right), \ w_j = w\left(x_j\right) \ \text{and} \ \eta_{i,j} = \eta\left(x_i, x_j\right)$$

represent discrete notations for functions g, β, w and η in (1.1), respectively.
The model equation (1.1) is approximated using the implicit finite difference scheme described in [4] and given below.

$$
\begin{cases}
\dfrac{u_j^{k+1}(q)-u_j^k(q)}{\Delta t}+\dfrac{g_j^k u_j^{k+1}(q)-g_{j-1}^k u_{j-1}^{k+1}(q)}{\Delta x}+w_j u_j^{k+1}(q) \\
+u_j^{k+1}(q)\displaystyle\sum_{l=1}^{N}\eta_{j,l}u_l^k(q)\ \Delta x=h_j^k(q) \\
g_0^k\ u_0^{k+1}(q)=\displaystyle\sum_{i=1}^{N}\beta_i^k u_i^{k+1}(q)\ \Delta x \\
P^{k+1}(q)=\displaystyle\sum_{j=1}^{N}u_j^{k+1}(q)\,\Delta x,
\end{cases}
\tag{2.2}
$$

where

$$h_j^k(q)=\frac{1}{2}\sum_{m=1}^{j}\eta_{j-m,m}u_{j-m}^k(q)\,u_m^k(q)\,\Delta x,$$

with initial condition

$$u_j^0=\frac{1}{\Delta x}\int_{(j-1)\Delta x}^{j\Delta x}u^0(x)dx \quad j=1,\cdots,N.$$

The ℓ^1, ℓ^∞ and BV norms are defined as

$$\left\|u^k(q)\right\|_1=\sum_{j=1}^{N}\left|u_j^k(q)\right|\Delta x, \qquad \left\|u^k(q)\right\|_\infty=\max_{j=0,\cdots,N}\left|u_j^k(q)\right|$$

and

$$\left\|u^k(q)\right\|_{BV}=\sum_{j=1}^{N}\left|D_h^-\left(u_j^k(q)\right)\right|\Delta x, \text{ respectively,}$$

where

$$D_h^-\left(u_j^k(q)\right)=\frac{u_j^k(q)-u_{j-1}^k(q)}{\Delta x},\ 1\le j\le N.$$

The following restrictions are also required for convergence of the above finite difference scheme (2.2):

(B_{mesh})

$$
\begin{cases}
\Delta x\dfrac{\beta_j^k}{g_0^k}+\dfrac{\Delta t}{\Delta x}g_j^k\left(1+\dfrac{\Delta t}{\Delta x}g_{j+1}^k+\Delta t w_{j+1}\right)^{-1}<1 & j=0,\cdots,N-1 \\
\Delta x\dfrac{\beta_N^k}{g_0^k}<1 & j=N.
\end{cases}
$$

(B_{other}) θ_1 and θ_2 are two constants satisfying

$$\sup_{(x,P)\in[0,x_{\max})\times[0,\infty)}\{\beta(x,P)-w(x)\}\le\theta_1$$

and

$$\sup_{(x,P)\in[0,x_{\max})\times[0,\infty)}\left|\frac{g(x+\delta,P)-g(x,P)}{\delta}+w(x)\right|\le\theta_2,$$

for sufficiently small $\delta>0$. Finally, $\max(\theta_1,\theta_2)\,\Delta t<1$.

The above implicit finite difference scheme was developed in [4, 5], where non-negativity and stability of this numerical method for suitably small Δt and Δx was proved. Further, the convergence of this scheme to a unique, bounded variation, weak solution of equation (1.1) was also shown. Hence the following result, concerning existence-uniqueness of the weak solution defined in equation (2.1) below, is recalled from [4].

Theorem 2.1. For any fixed $q \in Q$, equation (1.1) has a unique, bounded variation, weak solution.

Now the difference approximations $\left\{u_j^k(q)\right\}$ may be extended to a family of functions on $[0,T] \times [0, x_{\max}]$ by defining

$$U_{\Delta t, \Delta x}(t,x,q) = u_j^k(q), \quad (t,x) \in [t_{k-1}, t_k) \times [x_{j-1}, x_j), \ k = 1, \cdots, \Theta \text{ and } j = 1, \cdots, N.$$

The second step in solving the least squares minimization problem involves approximating the infinite dimensional parameter space Q by a sequence $\left\{Q^M\right\}$ of finite dimensional compact subsets of Q.

Thus, for computational purposes, one attempts to minimize the following approximate finite dimensional cost functional $J_{\Delta t, \Delta x}\left(q^M\right)$ defined as:

$$J_{\Delta t, \Delta x}\left(q^M\right) = \sum_{r=0}^{R} \left| \int_0^{x_{\max}} U_{\Delta t, \Delta x}(t_r, x, q^M) dx - \Pi_r \right|^2 \tag{2.3}$$

over Q^M, where $q^M \in Q^M$.

In the next section, the focus shall be on establishing convergence of minimizers of the approximated least squares problem (2.3) to a minimizer of the original least squares problem (1.2).

3 Convergence Results for Parameters

Investigation of convergence begins by proving that if there is a sequence of parameters $\left\{q^M\right\}$, where $q^M \in Q^M$, then the limit of the numerical solution of (2.2) corresponding to parameter q^M is the unique BV weak solution of (1.1) corresponding to parameter q^M.

Theorem 3.1. Let $U_{\Delta t, \Delta x}\left(t, x, q^M\right)$ denote solution of finite difference equation (2.2) corresponding to parameter q^M and let $u(t,x,q^M)$ be the unique, bounded variation, weak solution of equation (1.1), corresponding to parameter q^M. Then $U_{\Delta t, \Delta x}\left(t, \cdot, q^M\right) \to u(t, \cdot, q^M)$, in $L^1(0, x_{\max})$, uniformly for all $t \in [0,T]$, when $\Delta t, \ \Delta x \to 0$.

Proof: Following [3], define $u_j^{k,M} = u_j^k\left(q^M\right)$. Using assumptions on parameter space Q and proofs in [4], it can be seen that quantities $\left\|u^{k,M}\right\|_\infty$ and $\left\|u^{k,M}\right\|_{BV}$ are bounded, independent of M, Δt and Δx. Further, from [4], the approximants $\left\{u_j^{k,M}\right\}$ can be shown to satisfy the following Lipschitz condition in t:

$$\sum_{j=1}^{N} \left| \frac{u_j^{\alpha,M} - u_j^{\gamma,M}}{\Delta t} \right| \Delta x \le A_5 (\alpha - \gamma),$$

where A_5 is a fixed constant. Following procedures such as those given in [18], page 276, it can be shown that the approximants $\left\{u_j^{k,M}\right\}$ which represent the family of functions $U_{\Delta t, \Delta x}\left(t, x, q^M\right)$ on $[0,T] \times [0, x_{\max}]$, converge to a bounded variation function $\widetilde{u}\left(t, x, q^M\right)$ (along a subsequence) in $L^1(0, x_{\max})$, uniformly in t. Using uniqueness of this bounded variation solution stated in Theorem 2.1., it is only required to be shown that this solution $\widetilde{u}\left(t, x, q^M\right)$ is a weak solution of equation (1.1) corresponding to parameter q^M,

which will establish the result. To achieve this, multiply equation (2.2) corresponding to parameter q^M by $\phi_j^{k+1}\Delta t\Delta x = \phi\left(t_{k+1}, x_j\right)\Delta t\Delta x$ where $\phi \in C^1\left((0,T)\times(0,x_{\max})\right)$. Sum over indices $j = 1,\cdots,N$ and $k = 1,\cdots,\rho$ and use elementary algebraic manipulations to obtain the following for each $0 \le \rho \le \Theta$.

$$\begin{aligned}
&\sum_{j=1}^{N}\left(u_j^{\rho+1,M}\phi_j^{\rho+1} - u_j^{0,M}\phi_j^0\right)\Delta x - \sum_{k=1}^{\rho}\phi_0^{k+1}\sum_{j=1}^{N}\beta_j^{k,M}u_j^{k+1,M}\Delta x\Delta t \\
&-\sum_{k=1}^{\rho}\sum_{j=1}^{N}\left[u_j^{k,M}\left(\frac{\phi_j^{k+1}-\phi_j^k}{\Delta t}\right) + g_{j-1}^{k,M}u_{j-1}^{k+1,M}\left(\frac{\phi_j^{k+1}-\phi_{j-1}^{k+1}}{\Delta x}\right) - w_j^M u_j^{k+1,M}\phi_j^{k+1}\right]\Delta x\Delta t \qquad (3.1) \\
&= \sum_{k=1}^{\rho}\sum_{j=1}^{N}\phi_j^{k+1}\left[\frac{1}{2}\sum_{l=1}^{j}\eta_{j-l,l}^{M}u_{j-l}^{k,M}u_l^{k,M}\Delta x - u_j^{k+1,M}\sum_{l=1}^{N}\eta_{j,l}^{M}u_l^{k,M}\Delta x\right]\Delta x\Delta t.
\end{aligned}$$

Using techniques similar to those used in [11, 15, 18], it is seen that equation (3.1) converges to:

$$\begin{aligned}
&\int_0^{x_{\max}}\widetilde{u}\left(t,x,q^M\right)\phi(t,x)dx - \int_0^{x_{\max}}u^0(x)\phi(0,x)dx \\
&= \int_0^t\int_0^{x_{\max}}\widetilde{u}\left(\phi_s + g\phi_x - w\phi\right)dxds \\
&+\int_0^t\phi(s,0)\left(\int_0^{x_{\max}}\beta\left(x,P\left(s,q^M\right)\right)\widetilde{u}\left(s,x,q^M\right)dx\right)ds \\
&+\int_0^t\int_0^{x_{\max}}F\left(\widetilde{u}\left(s,x,q^M\right)\right)\phi dxds.
\end{aligned}$$

Hence the limit function $\widetilde{u}\left(t,x,q^M\right)$ is indeed the unique, bounded variation, weak solution of (1.1) corresponding to parameter q^M. This proves the theorem.

The following result follows from Theorem 3.1.

Corollary 3.2. $J_{\Delta t,\Delta x}\left(q^M\right) \to J\left(q^M\right)$ when $\Delta t, \Delta x \to 0$.

Proof: Follows from Theorem 3.1.

In the next theorem we show continuity of the approximate cost functional J in order to show the existence of a minimizer for the same.

Theorem 3.3. For fixed values of Δt and Δx, let $U_{\Delta t,\Delta x}(t,x,q^M)$ and $U_{\Delta t,\Delta x}(t,x,q)$ be solutions of the finite difference equation (2.2) corresponding to parameters q^M and q, respectively. Secondly, let $u(t,x,q^M)$ and $u(t,x,q)$ be unique, bounded variation, weak solutions of (1.1) corresponding to parameters q^M and q, respectively. Finally, let $q^M \to q$ in Q, when $M \to \infty$. Then

(a) $U_{\Delta t,\Delta x}(t,\cdot,q^M) \to U_{\Delta t,\Delta x}(t,\cdot,q)$ in $L^1(0,x_{\max})$, uniformly in $t \in [0,T]$.

(b) $u(t,\cdot,q^M) \to u(t,\cdot,q)$ in $L^1\left(0,x_{\max}\right)$, uniformly in $t \in [0,T]$, when $\Delta t, \Delta x \to 0$.

Proof: Let $\left\{u_j^{k,M}\right\}$ and $\left\{u_j^k\right\}$ denote the solutions $U_{\Delta t,\Delta x}(t,x,q^M)$ and $U_{\Delta t,\Delta x}(t,x,q)$, respectively, of the finite difference equation (2.2) corresponding to parameters q^M and q, respectively. Then, using techniques

similar to those used in lemma 16.6 of [18] and [4], let $v_j^{k,M} = u_j^{k,M} - u_j^k$ for $k = 0, \cdots, \Theta$ and $j = 0, \cdots, N$. The following system of difference equations hence follows:

$$\begin{aligned}
&\frac{v_j^{k+1,M} - v_j^{k,M}}{\Delta t} + D_h^- \left(g^M(x_j, P^{k,M}) u_j^{k+1,M} - g(x_j, P^k) u_j^{k+1} \right) + \left(w_j^M u_j^{k+1,M} - w_j u_j^{k+1} \right) \\
&= \frac{1}{2} \left(\sum_{l=1}^{j} \eta_{j-l,l}^M u_{j-l}^{k,M} u_l^{k,M} \Delta x - \sum_{l=1}^{j} \eta_{j-l,l} u_{j-l}^k u_l^k \Delta x \right) - \sum_{l=1}^{N} \eta_{j,l}^M u_j^{k+1,M} u_l^{k,M} \Delta x + \sum_{l=1}^{N} \eta_{j,l} u_j^{k+1} u_l^k \Delta x
\end{aligned} \tag{3.2}$$

with boundary condition

$$\begin{aligned}
g^M(0, P^{k,M}) u_0^{k+1,M} - g(0, P^k) u_0^{k+1} = \;& \sum_{i=1}^{N} \beta^M \left(x_i, P^{k,M} \right) v_i^{k+1,M} \Delta x \\
& + \sum_{i=1}^{N} \left(\beta^M(x_i, P^{k,M}) - \beta(x_i, P^k) \right) u_i^{k+1} \Delta x.
\end{aligned} \tag{3.3}$$

Multiply equation (3.2) by $\Delta x \text{ sgn} \left(v_j^{k+1,M} \right)$ and sum over indices $j = 1, \cdots, N$. Further, use notation w_j^M for $w^M(x_j)$, $\eta_{j,k}^M$ for $\eta^M(x_j, x_k)$ and so on to obtain

$$\begin{aligned}
\frac{\left\| v^{k+1,M} \right\|_1 - \left\| v^{k,M} \right\|_1}{\Delta t} \leq\;& -\sum_{j=1}^{N} D_h^- \left(g^M\left(x_j, P^{k,M}\right) u_j^{k+1,M} - g\left(x_j, P^k\right) u_j^{k+1} \right) \Delta x \text{ sgn } \left(v_j^{k+1,M} \right) \\
& + \sum_{j=1}^{N} \left(w_j u_j^{k+1} - w_j^M u_j^{k+1,M} \right) \Delta x \text{ sgn } \left(v_j^{k+1,M} \right) \\
& + \frac{1}{2} \sum_{j=1}^{N} \left(\sum_{l=1}^{j} \eta_{j-l,l}^M u_{j-l}^{k,M} u_l^{k,M} \Delta x - \sum_{l=1}^{j} \eta_{j-l,l} u_{j-l}^k u_l^k \Delta x \right) \Delta x \text{ sgn } \left(v_j^{k+1,M} \right) \\
& + \sum_{j=1}^{N} \left(-\sum_{l=1}^{N} \eta_{j,l}^M u_j^{k+1,M} u_l^{k,M} \Delta x + \sum_{l=1}^{N} \eta_{j,l} u_j^{k+1} u_l^k \Delta x \right) \Delta x \text{ sgn } \left(v_j^{k+1,M} \right) \\
& = I + II + III + IV.
\end{aligned}$$

Adding and subtracting terms leads to

$$\begin{aligned}
I = \;& -\sum_{j=1}^{N} D_h^- \left(g^M\left(x_j, P^{k,M}\right) v_j^{k+1,M} \right) \Delta x \text{ sgn} \left(v_j^{k+1,M} \right) \\
& + \sum_{j=1}^{N} D_h^- \left(\left[g^M\left(x_j, P^{k,M}\right) - g\left(x_j, P^k\right) \right] u_j^{k+1} \right) \Delta x \text{ sgn} \left(v_j^{k+1,M} \right) \\
& = i + ii.
\end{aligned}$$

Definition of D_h^- operator gives

$$\begin{aligned}
i \;&=\; g^M\left(0, P^{k,M}\right) v_0^{k+1,M} \text{ sgn} \left(v_1^{k+1,M} \right) - 2 \sum_{j \in Jump} g^M\left(x_j, P^{k,M}\right) \left| v_j^{k+1,M} \right| \\
&\leq\; g^M\left(0, P^{k,M}\right) v_0^{k+1,M} \text{sgn} \left(v_1^{k+1,M} \right) \leq \left| g^M\left(0, P^{k,M}\right) v_0^{k+1,M} \right|
\end{aligned}$$

where $Jump = \left\{1 \le j \le N-1 : v_j^{k+1,M} \cdot v_{j+1}^{k+1,M} < 0\right\}$.

Boundary condition equation (3.3) is used to obtain a bound on the term $\left|g^M\left(0,P^{k,M}\right)v_0^{k+1,M}\right|$.

$$
\begin{aligned}
g^M\left(0,P^{k,M}\right)v_0^{k+1,M} = \quad & \left(g\left(0,P^k\right) - g^M\left(0,P^{k,M}\right)\right)u_0^{k+1} + \sum_{i=1}^{N}\beta^M\left(x_i,P^{k,M}\right)v_i^{k+1,M}\Delta x \\
& + \sum_{i=1}^{N}\left(\beta^M\left(x_i,P^{k,M}\right) - \beta\left(x_i,P^k\right)\right)u_i^{k+1}\Delta x.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\left|g^M\left(0,P^{k,M}\right)v_0^{k+1,M}\right| \quad \le \quad & \left|g\left(0,P^k\right) - g^M\left(0,P^{k,M}\right)\right| \left\|u^{k+1}\right\|_\infty + \left\|\beta^M\right\|_\infty \left\|v^{k+1,M}\right\|_1 \\
& + \left|\beta^M\left(x_i,P^{k,M}\right) - \beta\left(x_i,P^k\right)\right| \left\|u^{k+1}\right\|_1 \\
\le \quad & \left|g\left(0,P^k\right) - g^M\left(0,P^k\right)\right| \left\|u^{k+1}\right\|_\infty + \left|g^M\left(0,P^k\right) - g^M\left(0,P^{k,M}\right)\right| \left\|u^{k+1}\right\|_\infty \\
& + \left\|\beta^M\right\|_\infty \left\|v^{k+1,M}\right\|_1 + \left|\beta^M\left(x_i,P^{k,M}\right) - \beta^M\left(x_i,P^k\right)\right| \left\|u^{k+1}\right\|_1 \\
& + \left|\beta^M\left(x_i,P^k\right) - \beta\left(x_i,P^k\right)\right| \left\|u^{k+1}\right\|_1 .
\end{aligned}
$$

Elementary algebra and summation by parts leads to

$$
\begin{aligned}
ii \quad = \quad & -\sum_{j=1}^{N} D_h^-\left(g\left(x_j,P^k\right) - g^M\left(x_j,P^{k,M}\right)\right)u_j^{k+1}\Delta x \ \mathrm{sgn}\left(v_j^{k+1,M}\right) \\
& -\sum_{j=1}^{N} D_h^-\left(u_j^{k+1}\right)\left(g\left(x_{j-1},P^k\right) - g^M\left(x_{j-1},P^{k,M}\right)\right)\Delta x \ \mathrm{sgn}\left(v_j^{k+1,M}\right) \\
\le \quad & \left|D_h^-\left(g\left(x_j,P^k\right) - g^M\left(x_j,P^{k,M}\right)\right)\right| \left\|u^{k+1}\right\|_1 \\
& + \left|g\left(x_{j-1},P^k\right) - g^M\left(x_{j-1},P^{k,M}\right)\right| \left\|u^{k+1}\right\|_{BV} \\
\le \quad & \left|D_h^-\left(g\left(x_j,P^k\right) - g^M\left(x_j,P^k\right)\right)\right| \left\|u^{k+1}\right\|_1 \\
& + \left|D_h^-\left(g^M\left(x_j,P^k\right) - g^M\left(x_j,P^{k,M}\right)\right)\right| \left\|u^{k+1}\right\|_1 \\
& + \left|g\left(x_{j-1},P^k\right) - g^M\left(x_{j-1},P^k\right)\right| \left\|u^{k+1}\right\|_{BV} \\
& + \left|g^M\left(x_{j-1},P^k\right) - g^M\left(x_{j-1},P^{k,M}\right)\right| \left\|u^{k+1}\right\|_{BV} .
\end{aligned}
$$

Adding and subtracting terms gives

$$
\begin{aligned}
II \quad = \quad & \sum_{j=1}^{N}\left(w_j - w_j^M\right)u_j^{k+1}\Delta x \ \mathrm{sgn}\left(v_j^{k+1,M}\right) - \sum_{j=1}^{N} w_j^M v_j^{k+1,M}\Delta x \ \mathrm{sgn}\left(v_j^{k+1,M}\right) \\
\le \quad & \left|w_j - w_j^M\right| \left\|u^{k+1}\right\|_1 + \left\|w^M\right\|_\infty \left\|v^{k+1,M}\right\|_1 .
\end{aligned}
$$

Changing order of summation leads to

$$
III = \frac{1}{2}\sum_{j=1}^{N} u_j^{k,M}\sum_{l=1}^{N}\eta_{j,l}^M u_l^{k,M}\Delta x\Delta x \ \mathrm{sgn}\left(v_j^{k+1,M}\right) - \frac{1}{2}\sum_{j=1}^{N} u_j^k\sum_{l=1}^{N}\eta_{j,l} u_l^k\Delta x\Delta x \ \mathrm{sgn}\left(v_j^{k+1,M}\right).
$$

Adding and subtracting terms,

$$
\begin{aligned}
III &= \frac{1}{2}\sum_{j=1}^{N} v_j^{k,M}\Delta x \sum_{l=1}^{N}\eta_{j,l}^{M} u_l^{k,M}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right) + \frac{1}{2}\sum_{j=1}^{N} u_j^{k}\Delta x\sum_{l=1}^{N}\eta_{j,l} v_l^{k,M}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right)\\
&\quad +\frac{1}{2}\sum_{j=1}^{N} u_j^{k}\Delta x\sum_{l=1}^{N}\left(\eta_{j,l}^{M}-\eta_{j,l}\right) u_l^{k,M}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right)\\
&\le \frac{1}{2}\left(\left\|\eta^{M}\right\|_\infty \left\|v^{k,M}\right\|_1 \left\|u^{k,M}\right\|_1 + \left\|u^{k}\right\|_1 \|\eta\|_\infty \left\|v^{k,M}\right\|_1 + \left\|u^{k}\right\|_1 \left\|u^{k,M}\right\|_1 \left|\eta_{j,l}^{M}-\eta_{j,l}\right|\right).
\end{aligned}
$$

Finally, using same techniques as in III, add and subtract terms and so on to get

$$
\begin{aligned}
IV &= -\sum_{j=1}^{N} u_j^{k+1,M}\Delta x\sum_{l=1}^{N}\eta_{j,l}^{M} u_l^{k,M}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right) + \sum_{j=1}^{N} u_j^{k+1}\Delta x\sum_{l=1}^{N}\eta_{j,l} u_l^{k}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right)\\
&= -\sum_{j=1}^{N} v_j^{k+1,M}\Delta x\sum_{l=1}^{N}\eta_{j,l}^{M} u_l^{k,M}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right) - \sum_{j=1}^{N} u_j^{k+1}\Delta x\sum_{l=1}^{N}\eta_{j,l} v_l^{k,M}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right)\\
&\quad +\sum_{j=1}^{N} u_j^{k+1}\Delta x\sum_{l=1}^{N}\left(\eta_{j,l}-\eta_{j,l}^{M}\right) u_l^{k,M}\Delta x \operatorname{sgn}\left(v_j^{k+1,M}\right)\\
&\le \left\|v^{k+1,M}\right\|_1 \left\|\eta^{M}\right\|_\infty \left\|u^{k,M}\right\|_1 + \left\|u^{k+1}\right\|_1 \|\eta\|_\infty \left\|v^{k,M}\right\|_1 + \left|\eta_{j,l}^{M}-\eta_{j,l}\right| \left\|u^{k+1}\right\|_1 \left\|u^{k,M}\right\|_1 .
\end{aligned}
$$

Gathering all the bounds above gives

$$
\begin{aligned}
\frac{\left\|v^{k+1,M}\right\|_1 - \left\|v^{k,M}\right\|_1}{\Delta t} &\le \left(\left|g\left(0,P^{k}\right)-g^{M}\left(0,P^{k}\right)\right| + \left|g^{M}\left(0,P^{k}\right)-g^{M}\left(0,P^{k,M}\right)\right|\right)\left\|u^{k+1}\right\|_\infty\\
&\quad + \left\|\beta^{M}\right\|_\infty \left\|v^{k+1,M}\right\|_1 + \left|\beta^{M}\left(x_i,P^{k,M}\right)-\beta^{M}\left(x_i,P^{k}\right)\right| \left\|u^{k+1}\right\|_1\\
&\quad + \left|\beta^{M}\left(x_i,P^{k}\right)-\beta\left(x_i,P^{k}\right)\right| \left\|u^{k+1}\right\|_1\\
&\quad + \left|D_h^{-}\left(g\left(x_j,P^{k}\right)-g^{M}\left(x_j,P^{k}\right)\right)\right| \left\|u^{k+1}\right\|_1\\
&\quad + \left|D_h^{-}\left(g^{M}\left(x_j,P^{k}\right)-g^{M}\left(x_j,P^{k,M}\right)\right)\right| \left\|u^{k+1}\right\|_1\\
&\quad + \left|g\left(x_{j-1},P^{k}\right)-g^{M}\left(x_{j-1},P^{k}\right)\right| \left\|u^{k+1}\right\|_{BV}\\
&\quad + \left|g^{M}\left(x_{j-1},P^{k}\right)-g^{M}\left(x_{j-1},P^{k,M}\right)\right| \left\|u^{k+1}\right\|_{BV}\\
&\quad + \left|w_j - w_j^{M}\right| \left\|u^{k+1}\right\|_1 + \left\|w^{M}\right\|_\infty \left\|v^{k+1,M}\right\|_1\\
&\quad + \frac{1}{2}\left\|\eta^{M}\right\|_\infty \left\|v^{k,M}\right\|_1 \left\|u^{k,M}\right\|_1 + \frac{1}{2}\left\|u^{k}\right\|_1 \|\eta\|_\infty \left\|v^{k,M}\right\|_1\\
&\quad + \left\|v^{k+1,M}\right\|_1 \left\|\eta^{M}\right\|_\infty \left\|u^{k,M}\right\|_1 + \frac{1}{2}\left\|u^{k}\right\|_1 \left\|u^{k,M}\right\|_1 \left|\eta_{j,l}^{M}-\eta_{j,l}\right|\\
&\quad + \left|\eta_{j,l}^{M}-\eta_{j,l}\right| \left\|u^{k+1}\right\|_1 \left\|u^{k,M}\right\|_1 + \left\|u^{k+1}\right\|_1 \|\eta\|_\infty \left\|v^{k,M}\right\|_1 .
\end{aligned}
$$

Since $q^M \to q$ in Q when $M \to \infty$, it follows that q^M is bounded. Further, since $\left\|u^{k,M}\right\|_1$ depends on boundedness of q^M (see [4]), $\left\|u^{k,M}\right\|_1$ is bounded as well. Using this and results stated earlier, it follows

that as $q^M \to q$ in Q, when $M \to \infty$ in the above inequality, there exist constants E and G such that

$$\frac{\left\|v^{k+1,M}\right\|_1 - \left\|v^{k,M}\right\|_1}{\Delta t} \leq E \left\|v^{k+1,M}\right\|_1 + G \left\|v^{k,M}\right\|_1 ,$$

which gives

$$\left\|v^{k,M}\right\|_1 \leq \left(\frac{1+G\Delta t}{1-E\Delta t}\right)^k \left\|v^{0,M}\right\|_1 .$$

This results in

$$\left\|U_{\Delta t,\Delta x}\left(t,\cdot,q^M\right) - U_{\Delta t,\Delta x}\left(t,\cdot,q\right)\right\|_1 \to 0$$

uniformly in $t \in [0,T]$. Hence the proof of (a). Further, letting $\Delta t, \Delta x \to 0$ and using Theorem 3.1. it can be seen that

$$\left\|u\left(t,\cdot,q^M\right) - u\left(t,\cdot,q\right)\right\|_1 \to 0$$

uniformly in $t \in [0,T]$. This proves (b). Hence the proof of the theorem.

The following result follows from Theorem 3.3.

Corollary 3.4. $J(q^M) \to J(q)$ as $q^M \to q$ in Q when $M \to \infty$.

Proof: Follows from Theorem 3.3.

Corollary 3.5. $J_{\Delta t,\Delta x}\left(q^M\right) \to J\left(q^M\right) \to J\left(q\right)$ when $\Delta t, \Delta x \to 0$ and $q^M \to q$ in Q when $M \to \infty$.

Proof: Follows from Corollaries 3.2. and 3.4.

In the next theorem, convergence result for solution of the least squares problem is established.

Theorem 3.6. Let Q be an infinite dimensional parameter space and $\left\{Q^M\right\}$ be a sequence of approximating finite dimensional compact subsets of Q. Further, if for each $q \in Q$ there exists $q^M \in Q^M$ such that $q^M \to q$ in Q when $M \to \infty$ and q^M is a minimizer of J over Q^M, then from compactness arguments, the sequence q^M has a subsequence q^{M_j} which converges to a minimizer $\overline{q}$ of J over Q.

Proof: Following an abstract least squares theory (see [9], pp. 143-145), proof follows from Corollary 3.5.

Numerical examples proving correctness of the theoretical results established above are presented next.

4 Numerical Examples

In this section, an example similar to that in [10], with final time $T = 5$ and maximum size $x_{\max} = 1$, is presented. The following known values for parameters g, β, w, η and initial condition u^0 are taken in the computations.

$$g(x,P) = 0.5(1-x)f(P), \;\; \beta(x,P) = 0.5xf(P), \;\; w(x) = 0.5,$$

$$\eta(x,y) = \begin{cases} 1-(x+y) & ; \quad x+y \leq 1 \\ 0 & ; \quad x+y > 1 \end{cases}$$

and

$$u^0(x) = \begin{cases} 10 & ; \quad 0 \le x \le 0.1 \\ 0 & ; \quad 0.1 < x \le 1. \end{cases}$$

$f(P)$, the only unknown function among the parameters to be approximated numerically, is taken as

$$f(P) = \exp(-4P)$$

in the first experiment and

$$f(P) = 0.1 + \frac{0.9}{1 + \exp\left[8\left(P - 0.5\right)\right]}$$

in the second experiment to computationally generate observed data. Model equation (1.1), for the parameter values given above, is solved numerically using the implicit finite difference scheme explained in [4] with mesh sizes fixed as $\Delta t = \Delta x = 0.01$. The observed data Π_r is then collected as follows:

$$\Pi_r = \int_0^{x_{\max}} u(t_r, x)dx, \; t_r = \frac{1}{5}r, \; r = 0, \cdots, 25.$$

Thus, for given fixed values of Υ and $P_0 > 0$, the parameter set Q is chosen as the D-closure of the following set:

$$\{f \in C_b\left[0, \infty\right) : \left|f(P)\right| \le \Upsilon, \; \left|f(P_1) - f(P_2)\right| \le \Upsilon \left|P_1 - P_2\right|, \; \forall P_1, P_2 \in [0, \infty)$$
$$\text{and } f(P) \text{ is constant for } P \ge P_{\max}, \text{ where } P_0 \ge P_{\max}\}.$$

A straightforward application of Arzelà-Ascoli theorem shows Q to be a compact subset of D. The infinite dimensional parameter space Q is approximated by the sequence $\left\{Q^M\right\}$ of finite dimensional compact subsets of Q, where

$$Q^M = \operatorname{span}\left\{\psi_0^M(P; P_{\max}), \psi_1^M(P; P_{\max}), \cdots, \psi_M^M(P; P_{\max})\right\},$$

M is a positive integer and $\left\{\psi_j^M(P; P_{\max})\right\}_{j=0}^M$ represents linear B-splines defined on the uniform partition

$$\left\{0, \frac{P_{\max}}{M}, \frac{2P_{\max}}{M}, \cdots, P_{\max}\right\}$$

of interval $[0, P_{\max}]$. The function $f(P) \in Q$ is approximated over Q^M in the following manner:

$$\left(I^M f\right)(P) = \sum_{j=0}^{M} f\left(j\frac{P_{\max}}{M}\right) \psi_j^M\left(P; P_{\max}\right) \text{ where } P \in [0, \infty).$$

The approximation of f on interval $[0, P_{\max}]$ is then extended to a continuous function on the non-negative real line via $\psi_j^M\left(P; P_{\max}\right) = \psi_j^M\left(P_{\max}; P_{\max}\right)$ for $P \ge P_{\max}$. Use of the Peano Kernel Theorem given in [17], gives the following result:

$$\lim_{M \to \infty} \left(I^M f\right) = f \text{ in } C_b\left[0, \infty\right), \text{ uniformly in } f, \text{ for } f \in Q.$$

Thus, if $f^M \in Q^M$ is given by

$$f^M(P) = \sum_{j=0}^{M} \zeta_j^M \psi_j^M\left(P; P_{\max}^M\right),$$

then the parameter estimation problem involves identifying $(M+2)$ coefficients $\left\{\zeta_j^M\right\}_{j=0}^M$ and $P_{\max}^M$ from a compact subset of R^{M+2}, in order to minimize the cost functional $J_{\Delta t, \Delta x}\left(q^M\right)$. In both experiments M is

chosen as 10. $\zeta_j^M = 0.5$, $j = 0, \cdots, M$ and $P_{\max}^M = 1$ are taken as initial guesses. Also, $P_0 = 2$ in both experiments.

For the least squares cost functional, the following penalized cost functional form is utilized:

$$J_\lambda(f) = \sum_{r=0}^{25} \left| \int_0^1 u(t_r, x, q)dx - \Pi_r \right|^2 + \lambda \int_0^1 |f'(P)|^2 \, dP.$$

The compactness of the embedding $H^1(0,1) \hookrightarrow L^2(0,1)$ enforces the compactness constraints. Properties of the regularized cost functional and the relationship between J_λ and J are discussed in [9, 12, 16]. The least squares problem is numerically solved using the *FORTRAN* routine *LMDIF1*, obtained from *NETLIB*, which uses the Levenberg-Marquardt algorithm.

In both experiments, data without noise and regularization parameter $\lambda = 10^{-6}$ are used initially. Comparison between exact and estimated function $f(P)$ is given in figures 1 and 2 for experiments 1 and 2, with $P_{\max}^M$ estimated as 0.92 and 0.99, respectively. These figures confirm theoretical results concerning convergence, proved earlier.

To check against error in measurement, noise with mean ($\mu = 0$) and standard deviation ($\sigma = 0.03$) is added to the computationally generated data Π_r. Exact and estimated functions for both experiments are shown together in figures 3 and 4, with estimates for $P_{\max}^M$ being 0.99 and 1.05, respectively. $\lambda = 3 \times 10^{-4}$ for noisy data in experiment 1 and $\lambda = 8 \times 10^{-4}$ for noisy data in experiment 2 were needed to obtain the best fit.

The least-squares cost functional values at the end of experiments 1 and 2 for data without noise were of the order of 10^{-8}, while the same for experiments 1 and 2 for data with noise had order 10^{-3}. Experiments 1 and 2, for data without noise, required between 15-20 hours of computation time on an *Ultra Sparc 2000* machine at the Computational Research Laboratory housed in the mathematics department at the University of Louisiana at Lafayette, Louisiana. Times for experiments 1 and 2, for data with noise, ranged from between 8 and 10 hours on the same computer system.

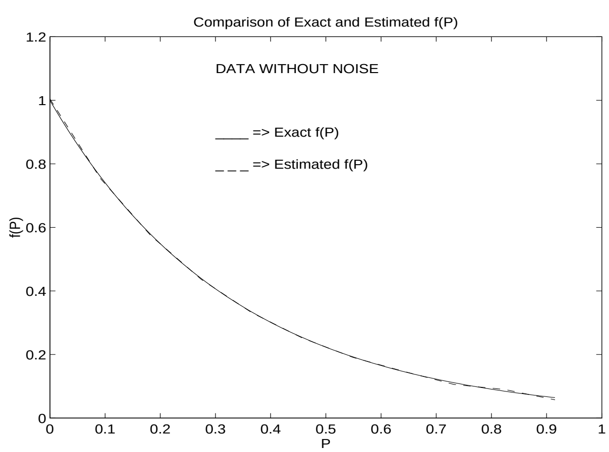


Figure 1: Exact $f(P) = e^{-4P}$

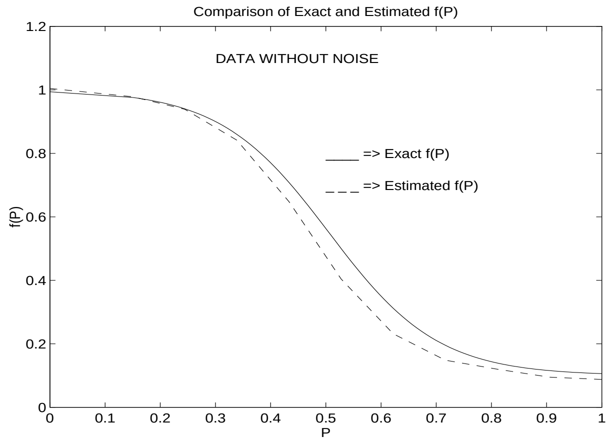


Figure 2: Exact $f(P) = 0.1 + 0.9/\left[1 + e^{8(P-0.5)}\right]$

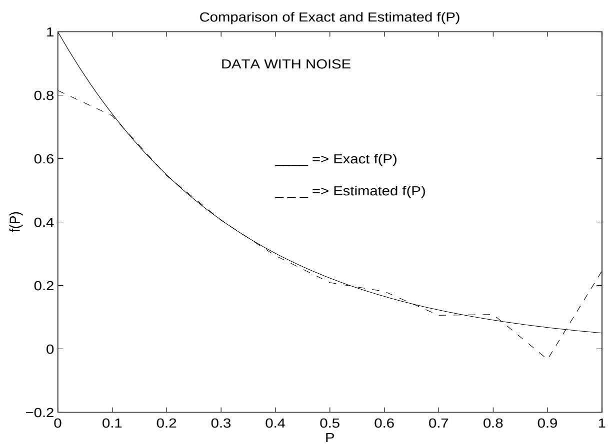


Figure 3: Exact $f(P) = e^{-4P}$

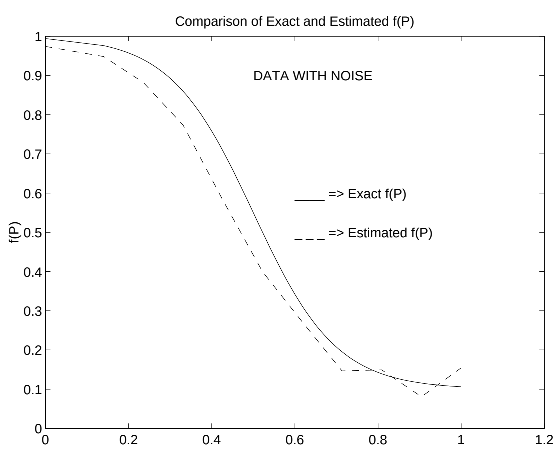


Figure 4: Exact $f(P) = 0.1 + 0.9/\left[1 + e^{8(P-0.5)}\right]$

5 Conclusion

Numerical results in this paper prove the feasibility of a least squares optimization method used to estimate parameters in a nonlinear, nonlocal, hyperbolic partial differential equation model containing a nonlinear reaction term. These results appear highly promising. Future efforts will involve parameter identification using experimentally observed data and a more sophisticated model than the one discussed in this paper. Estimating parameters in a two-dimensional reaction diffusion equation and in a nonlinear beam equation are also some other research issues to be possibly pursued by the author in the near future. All these results propose to appear in forthcoming papers to be submitted for possible publication.

References

1. A. S. Ackleh A. S. : Parameter Estimation in a Nonlinear Size-Structured Population Model, Adv. Sys. Sc. Appl., (1997), 315-320.

2. Ackleh A. S. : Parameter Estimation in a Structured Algal Coagulation Fragmentation Model, J. Nonlin. Anal. 28, (1997), 837-854.

3. Ackleh A. S. : Parameter Identification in Size-Structured Population Models with Nonlinear Individual Rates, Math. Comp. Model. 30, (1999), 81-92.

4. Ackleh A. S. and Ferdinand R. R. : A Finite Difference Approximation for a Nonlinear Size-Structured Phytoplankton Aggregation Model, Quart. J. Appl. Math. 57, (1999), 501-520.

5. Ackleh A. S. and Ito K. : An Implicit Finite Difference Scheme for the Nonlinear Size-Structured Population Model, Numer. Funct. Anal. Optimiz. 18, (1997), 865-884.

6. Banks H. T., Botsford L. W., Kappel F. and Wang C. : Estimation of Growth and Survival in Size-Structured Cohort Data: An Application to Larval Stiped Bass (Morone Saxatilis), J. Math. Biol. 30, (1991), 125-150.

7. Banks H. T., Botsford L. W., Kappel F. and Wang C. : Modeling and Estimation in Size Structured Population Models, T. G. Hallam, L. J. Gross, S. A. Levin (eds), Math Ecology, World Scientific, Singapore, 1988, pp. 521-541.

8. Banks H. T. and Fitzpatrick B. G. : Estimation of Growth Rate Distributions in Size Structured Population Models, Quart. J. Appl. Math. 49, (1991), 215-235.

9. Banks H. T. and Kunisch K. : Estimation Techniques for Distributed Parameter Systems, Birkhäuser, Boston, 1989.

10. Calsina A. and Saldana J. : A Model of Physiologically Structured Population Dynamics with a Nonlinear Growth Rate, J. Math. Biol. 33, (1995), 335-364.

11. Crandall M. G. and Majda A. : Monotone Difference Approximations for Scalar Conservation Laws, J. Math. Comp. 34, (1980), 1-21.

12. Engl H. W., Kunisch K. and Neubauwer A. : Tikhonov Regularization for the Solution of the Nonlinear Ill-Posed Problems I, Tech. Rep. 120, Technische Universität Graz, Graz, 1988.

13. Fitzpatrick B. G. : Parameter Estimation in Conservation Laws, J. Math. Syst. Est. Ctrl. 3, (1993), 413-425.

14. Fitzpatrick B. G. : Modeling and Estimation Problems for Structured Heterogeneous Populations, J. Math. Anal. App. 172, (1993), 73-91.

15. LeVeque R. J. : Numerical Methods for Conservation Laws, Birkhäuser, Boston, 1992.

16. Neubauer A. : Tikhonov Regularization for Nonlinear Ill-Posed Problems: Optimal Convergence Rates and Finite-Dimensional Approximation, Inv. Prob. 5, (1989), 541-557.

17. Schultz M. H. : Spline Analysis, Prentice-Hall, Englewood Cliffs NJ, 1973.

18. Smoller J. : Shock Waves and Reaction-Diffusion Equations, Springer-Verlag, Boston, 1994.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,281-288,
2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Operational and umbral methods for the solution of partial differential equations

Giuseppe Dattoli

ENEA – Unità di Fisica Teorica, Centro Ricerche Frascati, Via E. Fermi, 45 - 00044 - Frascati (RM), Italia - e-mail: dattoli@frascati.enea.it

Paolo E. Ricci

Università di Roma “La Sapienza”, Dipartimento di Matematica, P.le A. Moro, 2 00185 Roma, Italia - e-mail: riccip@uniroma1.it

Clemente Cesarano

ENEA – Unità di Fisica Teorica, Centro Ricerche Frascati, Via E. Fermi, 45 - 00044 - Frascati (RM), Italia & Ulm University, Department of Mathematics - Ulm, Germany - e-mail: cesarano@frascati.enea.it

Abstract

We employ operational and umbral caluculus methods to solve families of partial differential equations in a unified way. We show that the method can be extended to different forms of heat and D'Alembert equations providing explicit solutions hardly achieveble with conventional means.

2000 Mathematics Subject Classification. 44A45, 35A22, 35G10, 35K30, 39A99.
Key words and phrases. Operational methods, Integral transforms, Partial differential equations, Difference equations.

1 Introduction

In a previous paper [1] it has been shown that operational methods can be exploited to solve families of partial differential equations (p.d.e.), using a simple conceptual and computational effort.

To provide a specific example we will consider the following generalized D'Alembert problem

$$\begin{cases} \dfrac{\partial^2}{\partial t^2} f(x,t) = \widehat{\Omega}^2 f(x,t) \\ f(x,0) = g(x), \\ \dfrac{\partial}{\partial t} f(x,t)|_{t=0} = \gamma(x) \end{cases} \tag{1}$$

where $\widehat{\Omega}$ is a yet unspecified differential operator.

According to [1] we can solve (1) by treating $\widehat{\Omega}$ as an ordinary constant and the initial functions as integration constants, we can therefore write the general solution of equation (1) as

$$f(x,t) = \cosh(\widehat{\Omega}t)g(x) + \widehat{\Omega}^{-1}\sinh(\widehat{\Omega}t)\gamma(x) \tag{2}$$

where $\widehat{\Omega}^{-1}$ denotes the inverse of the operator $\widehat{\Omega}$.

In the case of $\widehat{\Omega} = \frac{\partial}{\partial x}$ we obtain D'Alembert solution, namely [1], [2]

$$\begin{aligned} f(x,t) &= \frac{1}{2}\left[g(x+t) + g(x-t)\right] + \frac{1}{2}\left[\phi(x+t) - \phi(x-t)\right], \\ \phi(x) &= \int_0^x \gamma(\xi)d\xi. \end{aligned} \tag{3}$$

The method of [1] is however general enough to obtain the solution for more general operators. In the following we will consider two different non trivial realizations of the operator $\widehat{\Omega}$ and use (2) to get explicit solutions.

In the case of $\widehat{\Omega} = x\frac{\partial}{\partial x}$ we apply equation (2) along with the rules [3]

$$\begin{aligned} e^{\lambda x \frac{\partial}{\partial x}} r(x) &= r(e^{\lambda}x), \\ \widehat{\Omega}^{-1} &= \int_0^{+\infty} e^{-s\widehat{\Omega}} ds \end{aligned} \tag{4}$$

thus getting

$$\begin{aligned} f(x,t) &= \frac{1}{2}\left[g(e^t x) + g(e^{-t}x)\right] + \\ &+ \frac{1}{2}\int_0^{+\infty}\left[\gamma(e^{t-s}x) - \gamma(e^{-t-s}x)\right] ds. \end{aligned} \tag{5}$$

The validity of the solution is limited to the case in which the integral on the r.h.s. of equation (5) converges.

In the case of $\widehat{\Omega} = x - \frac{\partial}{\partial x}$ we can apply the same procedure as before which along with the Weyl decoupling rule [3]

$$\begin{aligned} e^{\widehat{A}+\widehat{B}} &= e^{\widehat{A}}e^{\widehat{B}}e^{-\frac{k}{2}} \\ \text{if} \qquad \left[\widehat{A},\widehat{B}\right] &= \widehat{A}\widehat{B} - \widehat{B}\widehat{A} = k, \\ \text{and} \qquad \left[k,\widehat{A}\right] &= \left[k,\widehat{B}\right] = 0, \end{aligned} \tag{6}$$

yields

$$\begin{aligned} f(x,t) &= \frac{e^{-\frac{t^2}{2}}}{2}\left[e^{xt}g(x-t)+e^{-xt}g(x+t)\right]+ \\ &+ \frac{1}{2}\int_0^{+\infty} e^{-\frac{1}{2}(s^2+t^2)-sx}\left[e^{(x+s)t}\gamma(x-t+s)-e^{-(x+s)t}\gamma(x+t+s)\right]ds. \end{aligned} \tag{7}$$

The validity of the above solutions has been checked by means of a numerical integration in several cases.

The preliminary examples we have so far presented, yields an idea of the generality and usefulness of the method proposed in [1] which will be extended to other families of p.d.e. including umbral forms.

2 Operational background

To take some confidence, with the operational methods applied to the solutions of p.d.e. we consider the "heat" type equation

$$\begin{cases} \dfrac{\partial}{\partial t}\, f(x,t) = \left[q(x)\dfrac{\partial}{\partial x}\right]^2 f(x,t), \\ f(x,0) = g(x) \end{cases} \tag{8}$$

where $q(x)$ is a continous function non vanishing in the considered interval. The solution of (1) can be obtained by exploiting the method of the generalized Gauss transform [4]. By setting indeed

$$\widehat{\Gamma} = q(x)\frac{\partial}{\partial x} \tag{9}$$

we can write the formal solution of (7) as

$$f(x,t) = e^{t\,\widehat{\Gamma}^2} g(x). \tag{10}$$

By noting that

$$e^{t\,\widehat{\Gamma}^2} = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-\xi^2+2\sqrt{t}\,\xi\,\widehat{\Gamma}} d\xi \tag{11}$$

and by recalling that [3]

$$e^{\lambda q(x)\frac{\partial}{\partial x}} f(x) = f(F^{-1}(\lambda + F(x))), \tag{12}$$

where $F(x)$ is defined by

$$F(x) = \int_0^x \frac{d\xi}{q(\xi)}, \tag{13}$$

we find

$$f(x,t) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-\xi^2+2\sqrt{t}\xi q(x)\frac{\partial}{\partial x}} g(x)d\xi = \tag{14}$$
$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-\xi^2+2\sqrt{t}\xi} g(F^{-1}(2\xi\sqrt{t}+F(x)))d\xi.$$

In the case of $q(x)=x$ we have $F(x)=\ln(x)$, so that

$$f(x,t) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-\xi^2} g(e^{2\sqrt{t}\xi}x)d\xi. \tag{15}$$

This method can be applied to the solution of practical interest as the Black-Scholes equation often found in economical problems [4] (see Appendix).

Let us now consider the problem (8), with initial conditions

$$f(x,t)|_{x=0} = s(t), \qquad q(x)\frac{\partial}{\partial x}\, f(x,t)|_{x=0} = r(t). \tag{16}$$

The general solution can be found by exploiting the method outlined in the introductory section, thus finding

$$f(x,t) = \cosh\left(F(x)\left(\frac{\partial}{\partial t}\right)^{\frac{1}{2}}\right)s(t) + \widehat{\mathcal{D}}_t^{-\frac{1}{2}}\sinh\left(F(x)\left(\frac{\partial}{\partial t}\right)^{\frac{1}{2}}\right)r(t). \tag{17}$$

Where $\widehat{\mathcal{D}}_t^{-\frac{1}{2}}$ denotes the inverse of the half fractional derivative operator $(\partial/\partial t)^{\frac{1}{2}}$. The drawback of the above solution is the apparent necessity of dealing with fractional differential operators. Even though their use does not imply any particular problem, we note that they are not explicitly necessary, expanding indeed the hyperbolic functions in series we are left with the solution

$$f(x,t) = \sum_{n=0}^{\infty}\frac{F(x)^{2n}}{(2n)!}\left(\frac{\partial}{\partial t}\right)^n\left[s(t) + \frac{F(x)}{(2n+1)}\, r(t)\right], \tag{18}$$

whose evaluation does not require any use of fractional derivatives.

3 On partial differential, difference equations

After the discussion of the previous section, it is quite natural to consider an equation of the type

$$\begin{cases} \dfrac{\partial}{\partial t}\, f(x,t) = f(e^{\lambda}x,t) \\ f(x,0) = g(x). \end{cases} \tag{19}$$

We can easily recognize that the above equation can be solved using the so far outlined procedures, if we note that its r.h.s. can be cast in the form

$$f(e^{\lambda}x,t)=e^{\lambda x\frac{\partial}{\partial x}}f(x,t). \tag{20}$$

Accordingly we can write the solution of equation (19) as

$$f(x,t)=e^{te^{\lambda x\frac{\partial}{\partial x}}}\ g(x)=\sum_{n=0}^{\infty}\frac{t^n}{n!}\ g(e^{n\lambda}x). \tag{21}$$

More in general we can cast the solution of any equation of the type

$$\begin{cases}\dfrac{\partial}{\partial t}\ f(x,t)=f(F^{-1}(\lambda+F(x)),t),\\ f(x,0)=g(x)\end{cases} \tag{22}$$

in the form

$$f(x,t)=e^{te^{\lambda q(x)\frac{\partial}{\partial x}}}\ g(x)=\sum_{n=0}^{\infty}\frac{t^n}{n!}\ g(F^{-1}(n\lambda+F(x))), \tag{23}$$

and the validity of the solution is limited to the case in which the sum on the r.h.s. of equation (23) converges. This point will be more carefully discussed in the concluding section.

Equations of the above type are essentially partial finite difference equations and are generalizations of the case $(q(x)=1)$

$$\begin{cases}\dfrac{\partial}{\partial t}\ f(x,t)=f(x+\lambda,t)\\ f(x,0)=g(x),\end{cases} \tag{24}$$

whose solution is simply given by

$$f(x,t)=e^{te^{\lambda\frac{\partial}{\partial x}}}\ g(x)=\sum_{n=0}^{\infty}\frac{t^n}{n!}\ g(n\lambda+x). \tag{25}$$

Let us now consider a slightly more general problem, namely

$$\begin{cases}\dfrac{\partial^2}{\partial t^2}f(x,t)=f(e^{\lambda}x,t)\\ f(x,0)=g(x),\qquad \dfrac{\partial}{\partial t}f(x,t)|_{t=0}=\gamma(x).\end{cases} \tag{26}$$

in this case, by applying the same method as before, we find the solution in the form

$$f(x,t)=\sum_{n=0}^{\infty}\frac{t^{2n}}{(2n)!}\left[g(e^{2n\lambda}x)+\frac{t}{(2n+1)}\ \gamma(e^{(2n+1)\lambda}x)\right]. \tag{27}$$

Which can be easily generalized to the case with $q(x) \neq x$.

The results of this section show that the method we have developed are flexible enough to deal with different types of problems with a relatively modest computational effort.

4 Concluding remarks

The use of the so far developed method can be extended to less conventional forms of "Partial Differential Equations". To this aim we remind that in [1] it has been shown that equations of the type

$$\begin{cases} {}_l\widehat{\mathcal{D}}_x f(x,t) = \dfrac{\partial}{\partial t}\, f(x,t), \qquad {}_l\widehat{\mathcal{D}}_x := -\dfrac{\partial}{\partial x}\, x\, \dfrac{\partial}{\partial x} \\ f(0,t) = s(t) \end{cases} \tag{28}$$

where ${}_l\widehat{\mathcal{D}}_x$ is the Laguerre derivative [1]. Since the function

$$C_0(x) = \sum_{r=0}^{\infty} \frac{(-1)^r x^r}{(r!)^2} \tag{29}$$

is an eigenfunction of the Laguerre operator, we can write the solution of the problem (28) as

$$f(x,t) = C_0\left(x\frac{\partial}{\partial t}\right) s(t). \tag{30}$$

It is evident that the function (29) plays the same role of the exponential functions in the case of problems involving ordinary derivatives. More explicit solutions in terms of integral transform have been discussed in [1] and will not be reconsidered in the present paper.

Let us now go back to equation (1), written as

$$\widehat{\Gamma}^2 f(x,t) = \widehat{\Omega}^2 f(x,t) \tag{31}$$

where both $\widehat{\Gamma}$ and $\widehat{\Omega}$ are differential operators whose explicit realization has not been specified yet. By assuming that $\widehat{\Gamma}$ is a Laguerre derivative, i.e. that

$$\widehat{\Gamma} = -\frac{\partial}{\partial t}\, t\, \frac{\partial}{\partial t} \tag{32}$$

and that

$$\begin{aligned} &f(x,0) = g(x) \\ &\widehat{\Gamma} f(x,t)|_{t=0} = \gamma(x) \end{aligned} \tag{33}$$

we can solve our problem in the form given in equation (2), provided that cosh and sinh functions be replaced by

$$\begin{aligned} \cosh(\widehat{\Omega}) &= Ch(\widehat{\Omega}) \\ \sinh(\widehat{\Omega}) &= Sh(\widehat{\Omega}) \end{aligned} \tag{34}$$

where

$$\begin{aligned} Ch(\alpha) &:= \frac{C_0(\alpha) + C_0(-\alpha)}{2}, \\ Sh(\alpha) &:= \frac{C_0(\alpha) - C_0(-\alpha)}{2}, \end{aligned} \tag{35}$$

Which are defined in such a way that

$$\begin{aligned} {}_l\mathcal{D}_\alpha Ch(\alpha) &= Sh(\alpha) \\ {}_l\mathcal{D}_\alpha Sh(\alpha) &= Ch(\alpha). \end{aligned} \tag{36}$$

More explicit solutions, obtained by means of the integral Transform method introduced in [1] will be discussed in a forthcoming investigations.

Before closing this paper it is worth commenting on the validity of the solution (23). To be more clear we consider a specific example, namely $q(x) = x^2$, and accordingly [3]

$$\begin{aligned} e^{\lambda x^2 \frac{\partial}{\partial x}} f(x) &= f\left(\frac{x}{1-\lambda x}\right), \\ |\lambda x| &< 1. \end{aligned} \tag{37}$$

The sum in equation (23) should be limited to values of n such that $|n\lambda x| < 1$.

Further details will be discussed in a dedicated monography to these types of problems.

5 Appendix

In this appendix we discuss the solution of a problem of practical interest. We consider, indeed, the Black-Scholes equation, often occurring in economical problems, written in the form [4]

$$\frac{\partial}{\partial t}\, A = S^2 \frac{\partial^2}{\partial S^2}\, A + \lambda S \frac{\partial}{\partial S}\, A - \lambda A \tag{38}$$

which will be recast in a form more convenient for our purposes

$$\frac{\partial}{\partial t}\, A = \left(S\frac{\partial}{\partial S} + \frac{\lambda - 1}{2}\right)^2 A - \left(\frac{\lambda+1}{2}\right)^2 A. \tag{39}$$

Using the formalism discussed in this paper we can write the solution of equation (39) as

$$A(S,t) = e^{t\left[(S\frac{\partial}{\partial S}+\frac{\lambda-1}{2})^2-\left(\frac{\lambda+1}{2}\right)^2\right]}A(S,0) \tag{40}$$

thus getting according to the opeerational identities presented in the paper, we end up with

$$A(S,t) = \frac{e^{-\left(\frac{\lambda+1}{2}\right)^2 t}}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-\xi^2+(\lambda-1)\xi\sqrt{t}}A(e^{2\xi\sqrt{t}}S,0)d\xi. \tag{41}$$

The method discussed in the paper allows the extension of the solution to equations of the form

$$\frac{\partial^2}{\partial t^2}\,A = \left(S\frac{\partial}{\partial S}+\frac{\lambda-1}{2}\right)^2 A - \left(\frac{\lambda+1}{2}\right)^2 A \tag{42}$$

without any significant problem.

References

[1] G. Dattoli, I. Khomasuridze and P.E. Ricci, Operational methods, special polynomial and functions, and solution of partial differential equations, *Integral Transforms Spec. Funct.*, **15** (2004), 309–321.

[2] A.V. Bitzadze and D.F. Kalinichenko, *A Collection of Problems on Equations of Mathematical Physics*, MIR, Moscow, 1980.

[3] G. Dattoli, A. Torre, P.L. Ottaviani, L. Vázquez, Evolution operator equations: integration with algebraic and finite difference methods. Application to physical problems in classical and quantum mechanics, *Riv. Nuovo Cimento*, **2**, 1–133 (1997).

[4] H. Wyss, The fractional Blach-Scholes equation, *Fract. Calc. Appl. Anal.*, **3**, 51–61 (2000).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3,289-306,
2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Comparative analysis of image segmentation based on Euclidean metric and p-adic ultrametric

Jenny Benois-Pineau

* LABRI UMR CNRS 5800, University Bordeaux 1, 351,
Cours de la Liberation 33405 Talence, France
e-mail: jenny.benois@labri.fr
Andrei Khrennikov
International Center for Mathematical Modeling
in Physics and Cognitive Sciences,
MSI, University of Vaxjo, S-35195, Sweden
e-mail: Andrei.Khrennikov@vxu.se
Nikolai Kotovitch
Institute of System Analysis,
Russian Academy of Sciences, Moscow, Russia
e-mail: kotovich@niisi.msk.ru

Abstract

This paper presents an application of p-adic metric to image analysis in spectral domain. The main distinguishing feature of this metric is its ultrametricity. Clustering algorithm based on ultrametricity is developed. The algorithm produces partitions of the p-adic metric space having a very simple geometry: all clusters are represented as p-adic balls. p-adic clustering is used for segmentation of video and still images delivered in compressed format, e.g. MPEG1,2 for video and JPEG for still images. For some classes of images, the p-adic

clustering provides better visual quality. At the same time the p-adic algorithm ensures essential increasing of the computational speed.

AMS Subject Classification: 68T10, 68T30, 68W05, 11S99
Key words: ultrametric, clustering, DCT, MPEG – compressed.

1 Introduction

The problem of automatic analysis and interpretation of multimedia data in the compressed domain has got an importance since last few years. The reason for this is the increasing amount of multimedia archives and on-line services when the data such as video and still images are delivered in compressed format, e.g., MPEG1,2 [1] for video and JPEG [2] for still images. In particular, new user oriented services related to the broadcast, such as an automatic summarising and indexing of video content, realised on the receiver end in home multimedia devices have to deal with video data in MPEG2 [1].

Another application is related to the surveillance video. Here video is transmitted in compressed form and its analysis, such as object segmentation and tracking has also to be done in compressed form, see [3]. The problem of segmenting video or still images in the compressed domain has been already addressed in literature. In general case, when JPEG and MPEG1,2 still images are segmented, only the low frequency components the so called "DC images" of image spectrum are used [3]. New approaches have recently appeared [4] exploiting higher frequency spectral components in order to ensure a scalability of image (video) analysis algorithms. In this paper, we also propose to use more than low frequency component in DCT coefficients available in MPEG2 compressed video stream, thus exploiting textural characteristics in local areas in video frames. Particularity of our segmentation framework is the use of p-adic metric in clustering algorithms. The last one, having lexicographical properties implicitly introduces "an order of importance" of spectral components from low frequencies to higher. This corresponds to the well-known property of human visual system, that is its differential sensitivity to frequencies in image spectrum.

The p-adic metric is a so called ultrametric [5]. Instead of the usual triangle inequality it satisfies the "strong triangle inequality": in each triangle the third side is less than the maximum of two other sides. Thus the geometry

of an ultrametric space differs cruicially from the standard Euclidean geometry. One of the important (and rather unusual) features of this geometry is that if two balls have nonempty intersection, then one of them is a part of another. In fact, this property gives the possibility to provide clustering of images into collections of disjoint balls-clusters. The possibility of such a clustering is also related to the fact that in an ultrametric space so called chain distance coincides with the original distance. Thus a p-adic clustering of images has a simple geometry, namely a "ball-geometry".

During the last 10 years p-adic numbers are intensively applied to quantum physics, dynamical systems, spin glasses and cognitive science, see, e.g., [6], [7]. Cognitive applications [7] are closely related to our present investigation on image analysis. There are some evidences that cognitive systems (in particular, human brain) use p-adic (or more general ultrametric) clustering of information. The paper is organised as follows. Section 2 introduces fundamental concept of ultrametric spaces. In section 3, the overview of clustering methods is given and our new method of p-adic chain clustering is proposed. In Section 4 the method of choice of image vector data in spectral domain are given. Results and conclusion are presented in Section 5.

2 Ultrametric spaces

Abstract metric spaces are generalisations of the Euclidean space $\mathbf{R}^n = \{x = (\alpha_0, \ldots, \alpha_{n-1}) : x_j \in \mathbf{R}\}$ with the standard metric. However, in many applications (in particular, image analysis) we use spaces $X_{p,n}$ in which every point $x = (\alpha_0, \ldots, \alpha_{n-1})$ has discrete coordinates $\alpha_j = 0, \ldots, p-1$, where $p > 1$ is a natural number (α_j give the levels of discretisation of information). There are various possibilities to introduce a metric ρ on the space $X_{p,n}$. The standard choice of ρ is the Euclidean metric or some metric that is equivalent to the Euclidean one, e.g., Hamming metric.

In this paper we consider a so called p-adic metric, ρ_p, on the space $X_{p,n}$. The main distinguishing feature of this metric is its *ultrametricity*. This is an interesting special feature of the metric. Before considering the concrete p-adic ultrametric, it would be useful to discuss ultrametricity in the general topological framework. A metric ρ is said to be an *ultrametric* [5], if it satisfies the *strong triangle inequality*:

$$\rho(x, y) \leq \max \rho(x, z), \rho(z, y), x, y, z, \in X.$$

This inequality can be stated geometrically: *Each side of a triangle is at most as long as the longest one of the two other sides.* Let (X, ρ) be a metric space. We set $U_r(a) = \{x \in X : \rho(x, a) \leq r\}$ and $U_r^-(a) = \{x \in X : \rho(x, a) < r\}, r \in R_+, a \in X$. These are *balls* of radius r with center a. In the ultrametric space balls have the following properties [5]:

1. Each ball in X is both open and closed. Each point of a ball may serve as a centre. A ball may have infinitely many radii;

2. Let U and V be two balls in X. Then there are only two possibilities:
(a) balls are ordered by inclusion (i.e., $U \subset V$ or $V \subset U$);
(b) balls are disjoint.

We now introduce a so called p-adic ultrametric on the space $X_{p,n}$ of p-discrete vectors of the length n. Let $x = (\alpha_0, \ldots, \alpha_{n-1}), y = (\beta_0, \ldots, \beta_{n-1}) \in X_{p,n}$. We set $\rho(x, y) = 1/p^k$ if $\alpha_j = \beta_j, j = 0, 1, ..., k-1$, and $\alpha_k \neq \beta_k$. This is an ultrametric. Chain clustering algorithm that we address in this paper (see section 3) is, in fact, based on a so called "chain metric". We discuss this notion in the general metric framework.

A sequence of points $a = x_0, x_1, ..., x_{n-1}, x_n = b$ in a metric space (X, ρ) is called an ϵ-chain joining a and b if $\rho(x_k, x_{k+1}) \leq \epsilon$ for any $k \leq n$.

If there exists an ϵ-chain joining a and b they are said to be ϵ-linkable. A space (X, ρ) is Cantor connected if any two points can be joined by an ϵ-chain for any $\epsilon > 0$. The Euclidean space is Cantor connected. Ultrametric spaces are characterised by the following result [8]:

Theorem.
A metric space is ultrametric if and only if no two points $a \neq b$ in it are ϵ-linkable for any $\epsilon < \rho(a, b)$.

Let (X, ρ) be an arbitrary metric space. We set $\Delta(x, y) = \inf\{\epsilon > 0 : x \text{ and } y \text{ are } \epsilon\text{-linkable}\}$, $x, y \in X$. This function has all properties of an ultrametric, except non-degeneration (it can be that $\Delta(x, y) = 0$ for some $x \neq y$). It is a so called pseudo-ultrametric. It is called the *chain distance* between points x and y. We remark that Theorem implies that in an ultrametric space (X, ρ) the original metric ρ coincides with the corresponding chain distance. This topological fact simplifies clustering algorithms based on the computation of the chain distance [9] when they are modified for ultrametric (in particular, p-adic spaces).

3 Comparative analysis of image-clustering algorithms based on Euclidean and p-adic metrics

Cluster analysis (see [10]) is, in fact, a collection of various methods and algorithms of classification. Numerous attempts to classify methods of cluster analysis demonstrated that there exist hundreds of distinct classes (of methods). Such a variety is a consequence of a large number of possibilities to compute distances between various objects as well as distances between clusters in the process of clustering. We would like to pay attention to the following two groups of methods of cluster analysis: Agglomerative hierarchical algorithms (AH) and iterative clustering methods (IC).

3.1. Agglomerative hierarchical algorithms

The starting point in AH-methods is the consideration of all objects as separate (independent) clusters containing just a single element. The clustering procedure is based on (step by step) agglomeration of objects into clusters. The agglomeration process is based on a kind of a distance, ρ, between objects. The typical result of such a clustering is a hierarchical tree. There are many methods of agglomeration of objects into clusters, such as the method of the closest neighbour, the method of the maximally separated neighbours, unweighted pair-group, weighted pair-group method. All these methods are characterised by the fact that the distance between clusters is computed on the basis of distances between original objects in the set. We will base our investigation on the method of "chain clustering" described below.

3.1.1. Algorithm of chain clustering

The algorithm of chain clustering was proposed in [9]. This algorithm belongs to the group of AH-methods. Any object of the set of objects under the study can be chosen as the generator of clustering process. This object gets two labels, $n = 1$ (its number) and $\rho = 0$ (distances). Then we consider all other objects and take the object such that the distance ρ' between this object and the object with $n = 1$ is minimal. This new object gets $n = 2$ and $\rho = \rho'$. Then we consider all other objects and take the object such that the distance ρ'' between this object and the set of previously chosen objects is minimal. It is easy to see that thus defined, the distance ρ is the chain distance introduced in Section 2. At each step we take the object such

that the distance between this object and the set of objects that have been already labelled is minimal. This procedure is repeated until all objects are labelled. Finally, all objects are ordered (labelled) in a so called "chain" and each object in it has the label ρ - the distance to the set of previous objects. One of the ways to split the chain into clusters is based on the following procedure. Let $r_0 > 0$ be some constant (the parameter of clustering). We would like to build clusters in such a way that the chain distance between objects inside one cluster will be $\rho \leq r_0$ and at the same time the chain distance between objects belonging to different clusters will $\rho > r_0$. To do this, we consider ρ-labels of all objects and find objects such that $\rho > r_0$. Suppose that there are objects with numbers $n_1, \ldots, n_K$. The first cluster consists of all objects with $1 \leq n < n_1$; the second one - $n_1 \leq n < n_2$ and so on. By varying the parameter r_0 different partitions can be constructed. As we have already mentioned, in an ultrametric space the chain distance coincides with the original distance. Therefore chain-distance clustering is a clustering into balls in an ultrametric space. Thus in the ultrametric case from the beginning we could construct r_0-balls partition. The combinatorial complexity of such a method is $O(N^2)$, where N is the cardinal of initial vector set. It is the same precisely $N(N-1)/2$ in the worst and best cases.

3.1.2. Chain clustering with a fixed threshold

After the construction of the chain has been done, its splitting into clusters can be based on various criteria: fixed number of clusters, their statistical properties, etc. In the case of splitting of the chain with a fixed distance threshold, a simplified method for chain construction can be proposed that needs essentially less time. This method consists in the following.

Let us fix a distance between clusters, $r_0 > 0$. Any object can be taken as the initial object. It gets the label as belonging to the first cluster. The first cluster is constructed by grouping all objects such that the distance from them to the initial object is less than the threshold r_0. Then, for each of these new objects that were taken into the first cluster, the procedure is repeated (by considering only the objects not already belonging to the cluster). When there are no more objects left that could be collected into the first cluster, an object that does not belong to the first cluster is taken as the base of the second cluster and so on. If we have N objects, then in the worst case we need $N(N-1)/2$ operations of distance computation, but in the best case only N operations are needed.

In the case of the p-adic metric space, chain clustering with the fixed threshold works especially successfully. In this case the chain distance co-

incides with the ordinary (p-adic) metric. Therefore, in order to find the distance between an element and a cluster, it is sufficient to find the distance between this element and an arbitrary element of the cluster. So we need not compute distances to all elements of the cluster, as we have to do in the general case (and, in particular, the Euclidean case). The p-adic modification of the method of chain clustering with the fixed threshold is given by the following algorithm:

As the starting point we take any object. It gets the label of the first cluster. Then we take a new object. If the (p-adic) distance between the first and second objects is less than r_0, then the second object also gets the label of the first cluster; if not, then it gets the label of the second cluster. Then the whole process is repeated. The number of clusters increases. It is easy to see that here, for N objects and to obtain K clusters, we need not more than $N * K$ operations of the computation of distance.

3.2. Iteration methods

One of the most popular iteration methods of clustering is the method of K-means of MacQueen [11]. This method strongly differs from AH-methods. In opposite to AH-methods, here the user must fix number of clusters, K from the beginning. As in AH-methods, here we can choose different metrics as bases of clustering. Different algorithms of the method of K-means are also characterised by the way to choose centres of initiated clusters. One of the possible choices is given by so called Split-LBG process, that is generalized K-means approach which we implemented in this work. The method starts with $K = 2$ and optimises the partition by K-means approach. The way to add cluster centres proposed in [12, 13] consists in "splitting" of each already existing cluster centre $\mathbf{c_k}$ into two new centres $\mathbf{c}'_k = \mathbf{c}_k + \epsilon, \mathbf{c}''_k = \mathbf{c}_k - \epsilon$, where ϵ is a random vector of a weak energy . Thus, starting with 2 randomly chosen centres of clusters, the method allows to construct the partition of the vector space R^n into $K = 2^q$clusters by q steps of K-means optimisation.

4 Extraction image vectors from compressed video streams.

Let us now describe the nature of a vector space we deal with. The goal here is to perform image segmentation in a compressed domain of video standards

MPEG 1,2. The architecture of MPEG bit stream comprises three kinds of frames I, B, P. I-frames are coded entirely by DCT thus reducing spatial redundancy while P and B frames are motion-compensated. In this paper we consider the segmentation of I-frames. Here each block of 8x8 pixels in an image is coded by DCT, as in JPEG standard [2]. Video frames in MPEG2 are supposed to be represented in YUV color system, where Y is a grey-level (luminance) component and U and V are chrominance components. All three images Y,U,V are coded independently by JPEG (MPEG I-frame) coding algorithm. The coding method can be briefly described as follows. In each 8x8 pixel block, the original signal g(x,y) is centered :

$$f(x, y) = g(x, y) - g_0,$$

where $g_0 = 128$ for 8-bit depth.

Then a two dimensional DCT transform [1] is applied to obtain transform $F(u, v)$. The DCT coefficients are organized in blocks of 8x8 elements, where the coefficient $F(0, 0)$ called a "DC-coefficient" in the standard is 8 times the mean value of the original signal $f(x, y)$ in the block. Other coefficients, called "AC-coefficients" on the upper left corner in the block correspond to low frequencies in DCT spectrum and those in the right lower corner — to the high frequencies. The coefficients $F(u, v)$ are then quantized in order to reduce the information. The quantization in MPEG standards is applied in such a manner, that the coefficients at high frequencies are quantized more roughly than those at low frequencies. This is due to the fact that high frequencies in a signal spectrum correspond to noise and the human visual system is less sensitive to the loss of high frequencies in the decoded signal than to the low and medium frequencies.

A very important step in MPEG2 intra-frame encoding is the so called zig-zag scan. Here the spectral coefficients $F(u, v)$ for a block are registered in one-dimensional vector $QFS(n)$ with $n = 0, 1, \ldots, N \times N - 1$ according to the zig-zag order according to growing frequency. The data we use for segmentation are the de-quantized zig-zag ordered DCT coefficients of three components Y, U, V :

$$\mathbf{a} = (y_0, u_0, v_0, \ldots, y_{NxN-1}, u_{NxN-1}, v_{NxN_1}^T).$$

They can be truncated to $L - th$ co-ordinate, that is

$$\mathbf{a} = (a_0, \ldots, a_L)^T, \text{with } L = nc * 3 - 1, \tag{1}$$

where nc is the number of spectral coefficients retained. Thus the first three co-ordinates ($l = 0, 1, 2$) of vectors $\mathbf{a}_i$ characterise a "colour" in a block, as they represent the mean values of luminance and chrominance in the blocks, and further co-ordinates characterise the "texture" or "contours" inside blocks, as they correspond to signal variations inside a block. On the contrary to work [3] we use not only the low frequency component but also higher frequency coefficients characterising texture. In order to compute the p-adic distance and without being sensitive to insignificant variations, all vector coordinates were uniformly re-quantised. When clustering such vector spaces with fixed number of clusters, the question arises about the choice of this number. In literature good results are reported for natural images with methods similar to K-means (e.g. ISODATA) with the maximal number of clusters of 8. Nevertheless, if the problem is not only to segment image in principal regions but also to be able to reconstruct images based on cluster centres, then a larger number of clusters can be chosen based on visual satisfaction assessment.

Another important question is the choice of number of spectral DCT coefficients (that is the dimension of vectors (1)). We propose the following scheme of choice. According to the method to form the vectors (1) by zig-zag scanning of spectral coefficients in components Y,U,V the choice of $nc = 1$ coefficients in each component will supply a vector of mean values YUV (DC coefficients) in a block of 8x8 pixels in images. Incrementing the number of coefficients as $nc = 1, 3, 6, 10\ldots$ according to "slices" of zig-zag scan, the bands of high frequencies will be added to the vectors. As the original DCT coefficients have already undergone quantising and inverse quantising, the higher frequency coefficients equal to zero in a majority of cases. Therefore, the highest number of coefficients used in this work was $nc = 10$.

5 Results and discussion.

In order to prove the interest of the use of DCT coefficients for image segmentation, the discriminative power and efficiency of the p-adic metrics in clustering algorithms, the experiments were conducted on a large set of DCT-compressed images. Namely 250 I-frames of MPEG2 compressed movies "The time of lagoons", "A man from Tautavel", "Dulcimer Player", "Aquaculture in Mediterrannean Sea" SFRS® were processed. The goal of the first series of experiments performed was to show that for the same number

of clusters, the use of supplementary (AC) spectral coefficients better ensures the separation of clusters in natural images. Some results of segmentation by Split-LBG clustering with 8 clusters are given in Figure 1. The first column represents the DC images, obtained by replacing a block of 8x8 colour pixels in the original video frame by only one colour pixel representing the mean YUV vector of the block. The resolution of this images is therefore 90x72 pixels (based on CCIR601 initial resolution of frames). Then from left to right, the results of segmentation with progressively increasing number of spectral coefficients are shown. Images in columns 2-5 represent segmentation maps where the same colour corresponds to the same cluster in the vector space. These results observed on a whole data set show that using of AC spectral coefficients and not only DC, i.e., F(0,0) allows for better homogeneity of clusters and also for better separation of them. A typical example of such a better separation are the clusters corresponding to man's face and the background in images in the first range (see Figure 1.f, here the original zoomed DC image and segmentation maps with $nc = 1$ and $nc = 6$ respectively are shown from left to right). If we suppose that the same clustering method on colour images has to be performed in the original space and not in DCT domain, then for an image block of 8x8 pixels, the number of coefficients is 3x64=128. Thus the use of spectral coefficients also allows for a strong reduction of dimensionality of the vector space.

In the next series of experiments we compared an agglomerative hierarchical method (chain clustering) in the case of the Euclidean and the p-adic metrics. The results of this comparison are depicted in Figure 2. In Figure 2.a), odd rows starting from the first correspond to segmentation maps and the even rows represent the images reconstructed with DC coefficients. The odd columns correspond to the results for chain clustering with the Euclidean metric and the even rows depict results obtained for the p-adic metric. The comparative number of clusters is given in Figure 2.b) Analysing these results it can be stated that in order to produce segmentation of a good quality in the p-adic case we need less number of clusters than with the Euclidean metric. This is quite natural, because the first DCT coefficients play more important role in image formation than coefficients of higher orders. The computational time is also significantly lower for the p-adic case, as the computation of the p-adic distance is faster compared to the Euclidean distance (see Figure 2.c)).

Figure 3 depicts what we call the p-adic reconstruction of DC images. Here the DC images of "Dulciner Player" sequence are shown. They were

reconstructed after the p-adic chain clustering taking an arbitrary vector in each cluster as a centre of the cluster according to properties of an ultrametric. They have quite a natural aspect for relatively (compared to the Euclidean case) limited number of clusters.

Thus in this paper we proposed to use the p-adic metric for image segmentation in the spectral DCT domain and developed an adequate fast clustering method based on ultrametricity of the p-adic space. The results obtained on a large data set of natural images demonstrate the efficiency of this approach and its interest for segmentation of compressed (MPEG, JPEG) video and images with only a partial decompression of a bit-stream.

Acknowledgements

This paper was supported by the Swedish Royal Academy of Sciences and by the International Center for Mathematical Modelling of Växjö University grants.

References

1. M. Antonini, M. Barlaud and P. Mathieu, "Image coding using lattice vector quantisation of wavelet coefficients", *Proc. ICASSP '91,* paper M1.2, Toronto, 1991.

2. R. Engelking, General topology. *PWN,* Warsawa, 1977.

3. ISO/IEC 10918-1:1994 *Information technology – Digital compression and coding of continuous-tone still images* (JPEG).

4. ISO/IEC 13818-2:2000 ITU-T Rec. H.262 2nd Edition Z 75-004-2 4067400 *Generic coding of moving pictures and associated audio information: Video* (MPEG2 Video)

5. L. Kaufman, P. J. Rousseeuw *'Finding Groups in Data: An introduction to Cluster Analysis',* Wiley, 1990.

6. A. Khrennikov, *Non-Archimedean Analysis: Quantum Paradoxes, Dynamical Systems and Biological Models.* Kluwer Academic Publishers, Dordrecht, 1997.

7. J.-G. Kim, K.-W. Lee, J. Kim and H.-M. Kim 'Extraction of Moving Objects from MPEG-Compressed Video for Object-Based Indexing', *CBMI'99 European Workshop on Content – Based Multimedia Indexing,* Toulouse, France, 25-27 Oct. 1999

8. Y. Linde, A. Buzo, and R.M. Gray, 'An algorithm for vector quantiser design', *IEEE Transactions on communications,* vol. 28, 1980, pp. 84–95

9. J. MacQueen, 'Some methods for classification and analysis of multivariate observations', *Proc. Of the Fifth Berkley Symposium on Math. Stat. And Prob.,* pp. 281–296, 1967

10. W.H. Schikhof, Ultrametric calculus, *Cambridge University Press,* 1984.

11. E. Shepin, "SCRIT - optical character recognition program". Moscow, *MIAN,* 1992.

12. V.S. Vladimirov, I.V. Volovich, and E.I. Zelenov, *p-adic Analysis and Mathematical Physics,* World Scientific Publ., Singapore, 1994.

13. H.H. Yu ‘Scalable video browsing and searching via Q-metric’, *Pattern Recognition Letters,* Vol. 22 (5) (2001) pp. 493–502, 2001

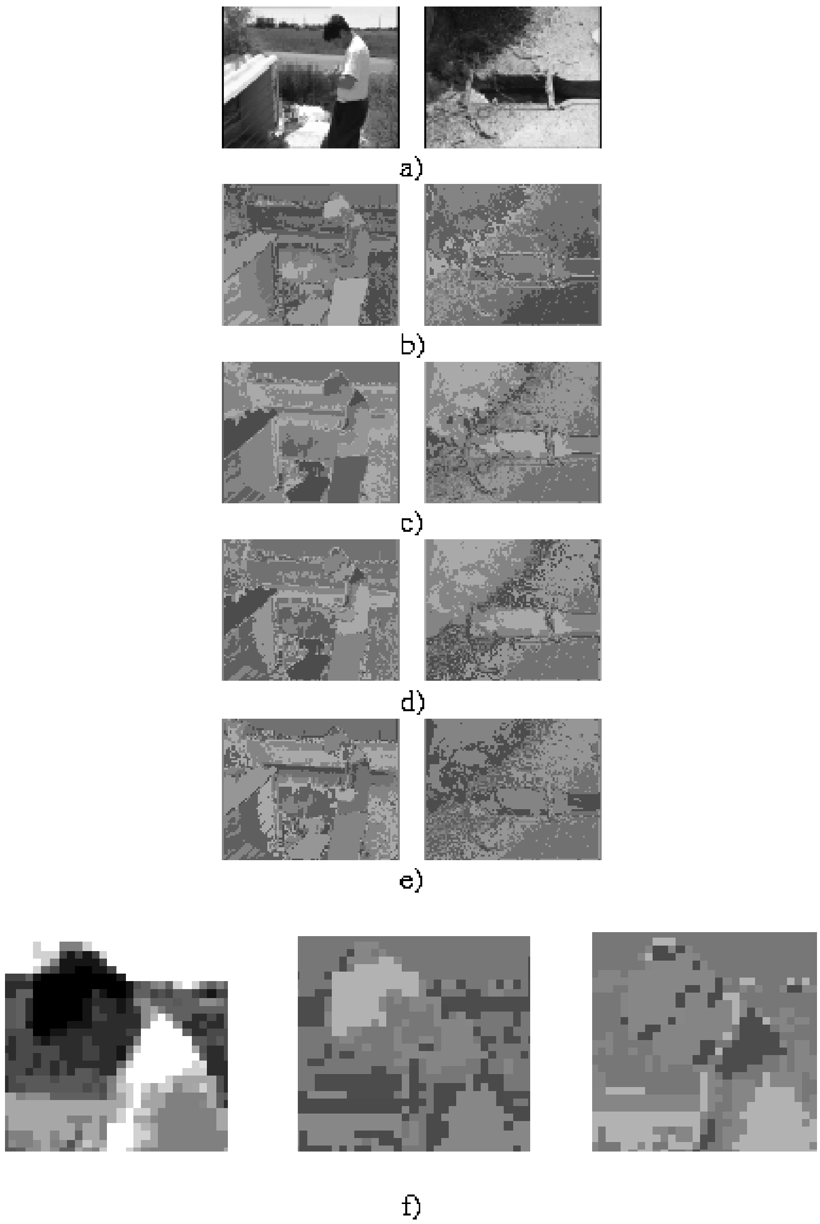


Figure 1: Figure 1. Results of clustering with Split-LBG method and Euclidean distance. From left to right : a) Original DC decoded image; b) Result of segmentation with nc = 1; c) nc=3; d) nc=6; e) nc=10; f) zoomed images a), b) and d).

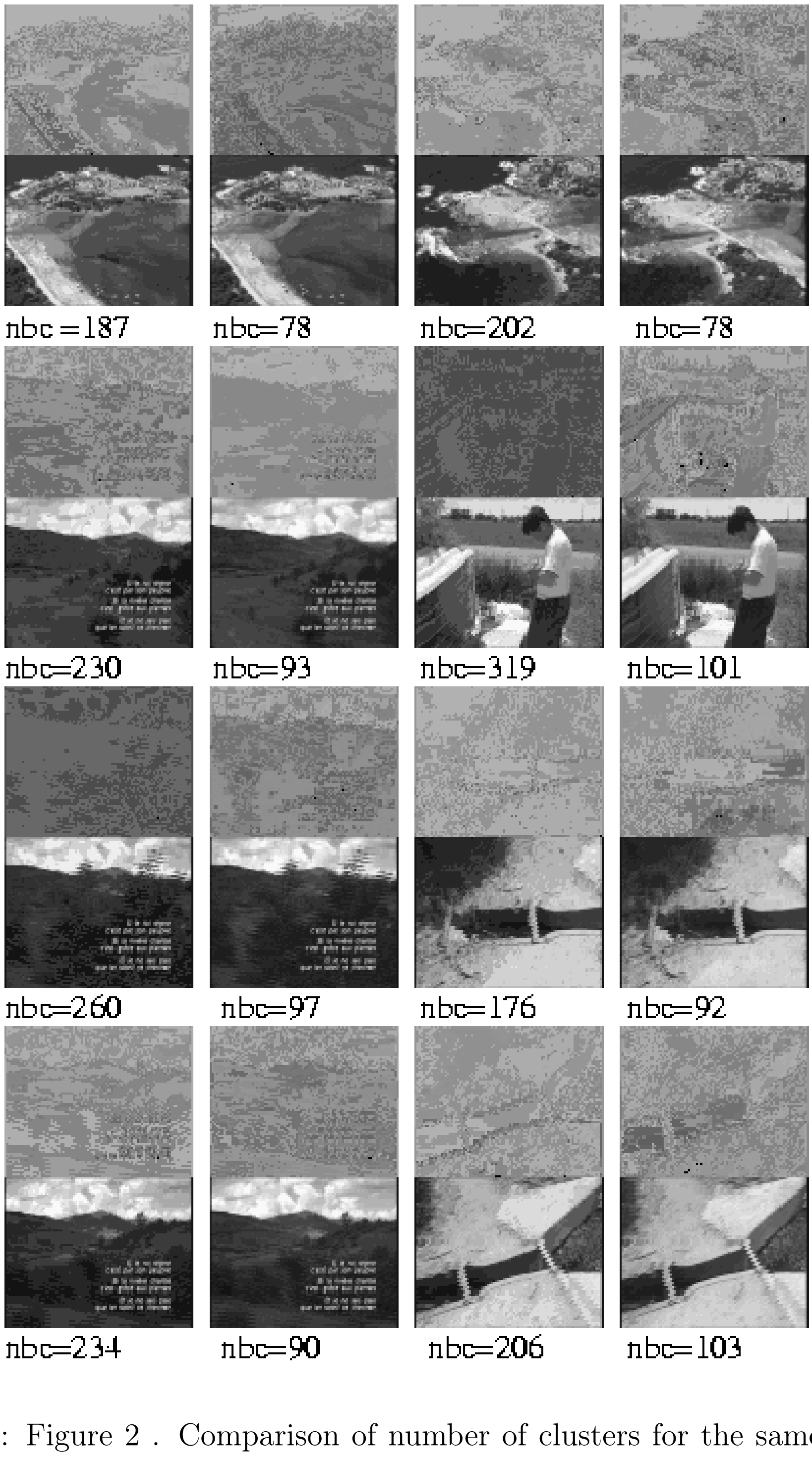
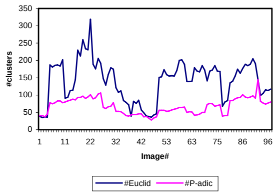
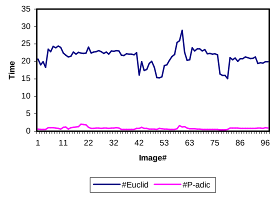


Figure 2: Figure 2 . Comparison of number of clusters for the same visual quality of reconstruction, NbrCoeff=5.Sequences "Aquaculture in Mediterrannean Sea", " The time of lagoons","A man from Tautavel", SFRS ®.



a)



b)

Figure 3: Figure 3. Computation parameters for data given at Figure 2: a) number of clusters, b) time of computation

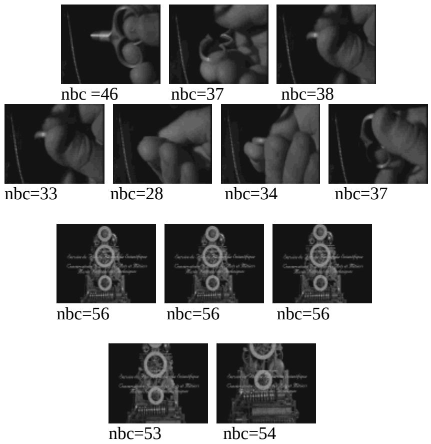


Figure 4: Figure 4. p-adic reconstruction of DC images. Sequence "Dulcimer Player", SFRS ®

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts hard copies in triplicate, and in English, should be submitted to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel.001. 901.678.3144
e-mail:ganastss@memphis.edu.

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.
Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.
2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (VISIT www.msci.memphis.edu/~ganastss/jcaam / to save a copy of the style file.)They should be carefully prepared in all respects. Submitted copies should be brightly printed (not dot-matrix), double spaced, in ten point type size, on one side high quality paper 8(1/2)x11 inch. Manuscripts should have generous margins on all sides and should not exceed 24 pages.
3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.
4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.
The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.
6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes. The main body of the paper should begin on page no. 1, if possible.
7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.
8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right,and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.
9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.
Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.
10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.

References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,
name of publication, volume number, inclusive pages, and year of publication.

Authors should follow these examples:

<u>Journal Article</u>

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

<u>Book</u>

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

<u>Contribution to a Book</u>

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.
12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set

off from the text. Use the appropriate superscript numeral for citation in the text.
13. After each revision is made please again submit three hard copies of the revised manuscript, including in the final one. And after a manuscript has been accepted for publication and with all revisions incorporated, manuscripts, including the TEX/LaTex source file and the PDF file, are to be submitted to the Editor's Office on a personal-computer disk, 3.5 inch size. Label the disk with clearly written identifying information and properly ship, such as:

Your name, title of article, kind of computer used, kind of software and version number, disk format and files names of article, as well as abbreviated journal name.

Package the disk in a disk mailer or protective cardboard. Make sure contents of disks are identical with the ones of final hard copies submitted!

Note: The Editor's Office cannot accept the disk without the accompanying matching hard copies of manuscript. No e-mail final submissions are allowed! The disk submission must be used.

14. Effective 1 Nov. 2005 the journal's page charges are $8.00 per PDF file page. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown in the Scope and Prices section.

No galleys will be sent and the contact author will receive one(1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

TABLE OF CONTENTS,JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.3.2004

VOLUME 2,NUMBER 4 OCTOBER 2004

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL
OF CONCRETE
AND APPLICABLE
MATHEMATICS

EUDOXUS PRESS,LLC

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send three copies of the contribution to the editor in-Chief typed in TEX, LATEX double spaced. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $250,Electronic $220,Print and Electronic $310.Individual:Print $77,Electronic $60,Print &Electronic $110.For any other part of the world add $25 more to the above prices for Print.

Single article PDF file for individual $8.Single issue in PDF form for individual $25.
The journal carries page charges $8 per page of the pdf file of an article,payable upon acceptance of the article within one month and before publication.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

Editorial Board
Associate Editors

Summability,Signal Analysis,Integral
Equations,
Measure Th.,Real Analysis

5) Francoise Bastin
Institute of Mathematics
University of Liege
4000 Liege
BELGIUM
f.bastin@ulg.ac.be
Functional Analysis,Wavelets

6) Paul L.Butzer
RWTH Aachen
Lehrstuhl A fur Mathematik
D-52056 Aachen
Germany
tel.0049/241/80-94627 office,
0049/241/72833 home,
fax 0049/241/80-92212
Butzer@rwth-aachen.de
Approximation Th.,Sampling Th.,Signals,
Semigroups of Operators,Fourier Analysis

7) Yeol Je Cho
Department of Mathematics Education
College of Education
Gyeongsang National University
Chinju 660-701
KOREA
tel.055-751-5673 Office,
055-755-3644 home,
fax 055-751-6117
yjcho@nongae.gsnu.ac.kr
Nonlinear operator Th.,Inequalities,
Geometry of Banach Spaces

8) Sever S.Dragomir
School of Communications and Informatics
Victoria University of Technology
PO Box 14428
Melbourne City M.C
Victoria 8001,Australia
tel 61 3 9688 4437,fax 61 3 9688 4050
sever.dragomir@vu.edu.au,
sever@sci.vu.edu.au
Math.Analysis,Inequalities,Approximation
Th.,
Numerical Analysis, Geometry of Banach
Spaces,
Information Th. and Coding

9) A.M.Fink
Department of Mathematics
Iowa State University
Ames,IA 50011-0001,USA

tel.352-392-9011
pardalos@ufl.edu
Optimization,Operations Research

24) Svetlozar T.Rachev
Dept.of Statistics and Applied Probability
Program
University of California,Santa Barbara
CA 93106-3110,USA
tel.805-893-4869
rachev@pstat.ucsb.edu
AND
Chair of Econometrics and Statistics
School of Economics and Business Engineering
University of Karlsruhe
Kollegium am Schloss,Bau II,20.12,R210
Postfach 6980,D-76128,Karlsruhe,Germany
tel.011-49-721-608-7535
rachev@lsoe.uni-karlsruhe.de
Mathematical and Empirical Finance,
Applied Probability, Statistics and Econometrics

25) Paolo Emilio Ricci
Universita' degli Studi di Roma "La Sapienza"
Dipartimento di Matematica-Istituto
"G.Castelnuovo"
P.le A.Moro,2-00185 Roma,ITALY
tel.++39 0649913201,fax ++39 0644701007
riccip@uniroma1.it,Paoloemilio.Ricci@uniroma1.it
Orthogonal Polynomials and Special functions,
Numerical Analysis, Transforms,Operational
Calculus,
Differential and Difference equations

26) Cecil C.Rousseau
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152,USA
tel.901-678-2490,fax 901-678-2480
ccrousse@memphis.edu
Combinatorics,Graph Th.,
Asymptotic Approximations,
Applications to Physics

27) Tomasz Rychlik
Institute of Mathematics
Polish Academy of Sciences
Chopina 12,87100 Torun, Poland
T.Rychlik@impan.gov.pl
Mathematical Statistics,Probabilistic
Inequalities

28) Bl. Sendov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria

tel.515-294-8150
fink@math.iastate.edu
Inequalities,Ordinary Differential Equations

10) Sorin Gal
Department of Mathematics
University of Oradea
Str.Armatei Romane 5
3700 Oradea,Romania
galso@uoradea.ro
Approximation Th.,Fuzzyness,Complex Analysis

11) Jerome A.Goldstein
Department of Mathematical Sciences
The University of Memphis,
Memphis,TN 38152,USA
tel.901-678-2484
jgoldste@memphis.edu
Partial Differential Equations,
Semigroups of Operators

12) Heiner H.Gonska
Department of Mathematics
University of Duisburg
Duisburg,D-47048
Germany
tel.0049-203-379-3542 office
gonska@informatik.uni-duisburg.de
Approximation Th.,Computer Aided Geometric Design

13) Dmitry Khavinson
Department of Mathematical Sciences
University of Arkansas
Fayetteville,AR 72701,USA
tel.(479)575-6331,fax(479)575-8630
dmitry@uark.edu
Potential Th.,Complex Analysis,Holomorphic PDE,Approximation Th.,Function Th.

14) Virginia S.Kiryakova
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms,
Fractional Calculus

15) Hans-Bernd Knoop
Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676
bsendov@bas.bg
Approximation Th.,Geometry of Polynomials,
Image Compression

29) Igor Shevchuk
Faculty of Mathematics and Mechanics
National Taras Shevchenko
University of Kyiv
252017 Kyiv
UKRAINE
shevchuk@univ.kiev.ua
Approximation Theory

30) H.M.Srivastava
Department of Mathematics and Statistics
University of Victoria
Victoria,British Columbia V8W 3P4
Canada
tel.250-721-7455 office,250-477-6960 home,
fax 250-721-8962
harimsri@math.uvic.ca
Real and Complex Analysis,Fractional Calculus and Appl.,
Integral Equations and Transforms,Higher Transcendental
Functions and Appl.,q-Series and q-Polynomials,
Analytic Number Th.

31) Ferenc Szidarovszky
Dept.Systems and Industrial Engineering
The University of Arizona
Engineering Building,111
PO.Box 210020
Tucson,AZ 85721-0020,USA
szidar@sie.arizona.edu
Numerical Methods,Game Th.,Dynamic Systems,
Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources
Management

32) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
Approximation Theory

33) Manfred Tasche
Department of Mathematics
University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,

knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

18) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

Image Processing,Harmonic Analysis

34) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

35) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.4,307-325,2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

A McShane integral for multifunctions

A. Boccuto - A. R. Sambucini

Department of Mathematics and Informatics,
via Vanvitelli, 1 - I-06123 Perugia (ITALY)
e-mail:boccuto@dipmat.unipg.it, matears1@unipg.it

Abstract

In this paper we introduce a "generalized" McShane integration for Banach-valued multifunctions with weakly compact and convex values and we give also a comparison between this integration and the Aumann integration.

1991 AMS Mathematics Subject Classification: 28B05, 28B20, 26E25, 46B20, 54C60
Key words: Aumann integral, McShane integral, multifunctions, Banach spaces, Rådstrom embedding theorem.

1. Introduction

The notion of integral of a multivalued function is very useful in many branches of mathematics like mathematical economics, control theory, differential inclusions, convex analysis, etc. It has been introduced by

many authors and in different ways. The first was Aumann in 1965, in order to apply it to general equilibria in economics. This integral was built using selections, but some properties were missing, so Debreu introduced the multivalued Bochner integral. In both cases the definition of measurable multifunction is crucial since it is necessary to ensure that at least a selection exists. Many authors worked on the problem of measurability of multifunctions; we quote here for example [4, 13, 11, 9, 10, 2] for the countably additive case and [20] for a review in the finitely additive case.

Here we introduce a new kind of multivalued integral which does not need a priori the notion of measurability; this fact looks interesting for example in differential inclusions. The idea comes out from a discussion with Prof. Jan Andres during a congress in 2000 and was presented in 2003 at the XVII Congress of U.M.I..

Our starting point is a paper by Jarník and Kurzweil [14] in which the authors proposed a new definition based on Kurzweil-Henstock "selections" for $\mathbb{R}^n$-valued multifunctions, defined in a bounded interval of $\mathbb{R}$. Jarník and Kurzweil applied it to differential inclusions and showed that under suitable conditions (namely compactness of values) this integral coincides with the Aumann's one.

Here we extend these results in two directions: we consider in fact multifunctions defined in the whole real line and moreover taking values in a Banach space not necessarily separable. In particular in section 3 we introduce the $(\star)$-integral by using McShane integrable single valued functions and then we compare it with the Aumann integral. Finally, in section 4, making use of the Rådstrom embedding theorem, the McShane multivalued integral is introduced and compared with the $(\star)$ and Aumann integrals. When the McShane multivalued integral exists, then the $(\star)$-integral exists too and it coincides with it, and so all the properties

of the single valued McShane integral are inherited by the multivalued one.

2. Preliminaries and known results on the generalized McShane integral.

The generalized McShane integral (McShane integral briefly), as a limit of suitable Riemann sums, was developed in the vector valued case by Fremlin in [7]. In this section, we assume that S is a space and $\mathcal{T}$ a topology on S making $(S, \mathcal{T}, \Sigma, \mu)$ a non-empty σ-finite quasi-Radon measure space which is *outer regular,* namely such that

$$\mu(B) = \inf\{\mu(G) : B \subseteq G \in \mathcal{T}\} \quad \forall\, B \in \Sigma.$$

A *generalized McShane partition* P of S ([7, Definitions 1A]) is a disjoint sequence $(E_i, t_i)_{i\in\mathbb{N}}$ of measurable sets of finite measure, with $t_i \in S$ for every $i \in \mathbb{N}$ and $\mu(S \setminus \bigcup_i E_i) = 0$.

A *gauge* on S is a function $\Delta : S \to \mathcal{T}$ such that $s \in \Delta(s)$ for every $s \in S$. A generalized McShane partition $(E_i, t_i)_i$ is *subordinate to* Δ if $E_i \subset \Delta(t_i)$ for every $i \in \mathbb{N}$.

From now on with the symbol $\mathcal{P}$ we denote the class of all generalized McShane partitions of $[a, b]$, and with $\mathcal{P}_\Delta$ those elements from $\mathcal{P}$ that are subordinate to Δ.

Let X be a Banach space. We say that:

Definition 1 [7, Definitions 1A] A function $f : S \to X$ is *McShane integrable,* with integral w, if for every $\varepsilon > 0$ there exists a gauge $\Delta : S \to \mathcal{T}$ such that

$$\limsup_{n\to+\infty} \left\| w - \sum_{i=1}^{n} \mu(E_i) f(t_i) \right\| \le \varepsilon$$

for every generalized $\mathcal{P}_\Delta$ McShane partition $(E_i, t_i)_i$. In this case, we write $\int_S f = w$.

For the properties of the McShane generalized integral we suggest the quoted article [7] by Fremlin. Here we recall only this result which will be used later:

[7, Lemma 1J] Let $f : S \to X$ be a function. Then, for every $\varepsilon > 0$, there exists a gauge $\Delta : S \to \mathcal{T}$ such that

$$\sum_{i=1}^{\infty} \mu(E_i) \|f(t_i)\| \leq \overline{\int}_S \|f(t)\| \mu(dt) + \varepsilon,$$

whenever $(E_i, t_i)_i$ is a generalized $\mathcal{P}_\Delta$ McShane partition of S and $\overline{\int}_S \|f(t)\| \mu(dt)$ denotes outer integration, namely

$$\overline{\int}_S \|f(t)\| \mu(dt) := \inf \left\{ \int_S g(t)\mu(dt), g \in L^1(\mathbb{R}), \|f(t)\| \leq g(t) \right\}.$$

Fremlin in [7] studied also the relationship among this integral and the usual "strong" and "weak" integrals in Banach spaces. In particular this new integral, which coincides with the classical one in $\mathbb{R}$, is weaker than the Bochner and stronger than the Pettis one. In fact Bochner integrability implies McShane integrability and the two integrals agree ([7, Theorem 1K]), while McShane integrability implies Pettis integrability and the two integrals agree ([7, Theorem 1Q]). Moreover, if the Banach space X is separable, then McShane and Pettis integrability coincide ([7, Corollary 4C]).

3. Applications to multivalued integration

Here we introduce a new kind of multivalued integral. There are in the literature several papers on Aumann integration and other multivalued integrations; see for example [1], [20] and their bibliography. Note that, in all existing multivalued integration theories, in order to define the multivalued integrals, a notion of measurability or "total measurability"

is required. For the kind of integrability that we will introduce, no measurability is required a priori and so we can define a multivalued integral also in non separable Banach spaces.

Throughout this section, let $S = [a,b]$, where $a, b \in [-\infty, +\infty]$, $a < b$. Moreover, assume that $\mathcal{T}$, Σ and μ are respectively the families of all open subsets of $[a,b]$, the σ-algebra of all Lebesgue measurable subsets of $[a,b]$ and the Lebesgue measure on $[a,b]$ respectively.

Let $cwk(X)$ $[ck(X)]$ denote the family of all convex and weakly compact [respectively convex and compact] subsets of a Banach space X. We denote with the symbol $d(x, C)$ the usual distance between a point and a nonempty set $C \subset X$, namely $d(x,C) = \inf\{\|x-y\| : y \in C\}$, and by $\mathcal{U}(C,\varepsilon)$ the ε-neighborhood of the set C, i.e.

$$\mathcal{U}(C,\varepsilon) = \{z \in E : \exists\, x \in C \text{ with } \|x - z\| \leq \varepsilon\}.$$

Observe that, if C is convex, then $\mathcal{U}(C,\varepsilon) = \mathrm{co}(\mathcal{U}(C,\varepsilon))$.
If C, D are two nonempty subsets of X, we denote with the symbol $e(C,D)$ the excess of C with respect to D, namely $e(C,D) = \sup\{d(x,D) : x \in C\}$, while the Hausdorff distance between C and D is $h(C,D) = \max\{e(C,D), e(D,C)\}$. We remember that $h(C,D) = 0$ if and only if $cl\{C\} = cl\{D\}$, where the symbol $cl\{\cdot\}$ denotes the closure of the considered set with respect to the norm topology.

Like in [14] we define a multivalued integral in the following way:

Definition 2 Let $F : [a,b] \to 2^X \setminus \emptyset$ be a multifunction. We call $(*)$-*integral* of F over $[a,b]$ the set $\Phi(F,[a,b])$ given by:

$$\begin{aligned}\Phi(F,[a,b]) \;=\; &\{x \in X : \forall\, \varepsilon > 0, \exists \text{ a gauge } \Delta : \text{ for every generalized}\\ &\mathcal{P}_\Delta \text{ McShane partition} (E_i, t_i)_{i\in\mathbb{N}} \text{ there holds}\\ &\limsup_n \, d(x, \sum_{i=1}^n F(t_i)\mu(E_i)) \leq \varepsilon\},\end{aligned}$$

where, as usual, $\sum_{i=1}^n F(t_i)\mu(E_i) := \{\sum_{i=1}^n x_i\mu(E_i) : \; x_i \in F(t_i)\}$.

Observe that, if F is single-valued, then $\Phi(F,[a,b])$ coincides with the McShane integral, if it exists. We now show that:

Proposition 1 *If F is bounded valued, then*

$$\Phi(F,[a,b]) = \bigcap_{\varepsilon>0} \bigcup_{\Delta} \bigcap_{(E_i,t_i)_i \in \mathcal{P}_\Delta} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon). \qquad (1)$$

Proof. Let $z \in \Phi(F,[a,b])$; for every $\varepsilon > 0$, there exists a gauge $\Delta(\varepsilon/2)$ such that for every generalized $\mathcal{P}_\Delta$ McShane partition $(E_i,t_i)_i$

$$\limsup_n \, d(z, \sum_{i=1}^{n} F(t_i)\mu(E_i)) = \inf_{m\geq 1} \sup_{n\geq m} d(z, \sum_{i=1}^{n} F(t_i)\mu(E_i)) \leq \varepsilon/2.$$

From this it follows that there exists $m \in \mathbb{N}$ such that

$$d(z, \sum_{i=1}^{n} F(t_i)\mu(E_i)) \leq \varepsilon \quad \text{ for every } n \geq m,$$

and thus

$$z \in \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon).$$

Hence, $z \in \bigcap_{\varepsilon>0} \bigcup_{\Delta} \bigcap_{(E_i,t_i)_i \in \mathcal{P}_\Delta} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon)$.

Conversely, let $z \in \bigcap_{\varepsilon>0} \bigcup_{\Delta} \bigcap_{(E_i,t_i)_i \in \mathcal{P}_\Delta} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon)$. Then, for every $\varepsilon > 0$, there exists a gauge Δ such that, for every generalized $\mathcal{P}_\Delta$ McShane partition $(E_i,t_i)_i$,

$$z \in \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon),$$

which means that for every $\varepsilon > 0$, there exists a gauge Δ such that, for every generalized $\mathcal{P}_\Delta$ McShane partition $(E_i,t_i)_i$,

$$\limsup_n \, d(z, \sum_{i=1}^{n} F(t_i)\mu(E_i)) \leq \varepsilon,$$

namely $z \in \Phi(F,[a,b])$. $\square$

Remark 1 **(a)** Observe that, by definition, the set $\Phi(F,[a,b])$ is closed; in fact if $(z_n)_n$ is a sequence in $\Phi(F,[a,b])$ which converges to $z \in X$ then, for every $\varepsilon > 0$ there exist an integer k and a gauge Δ_k such that for every generalized $\mathcal{P}_{\Delta_k}$ McShane partition $(E_i,t_i)_i$

$$\|z - z_k\| \le \varepsilon/2, \quad \limsup_{n\to\infty} d(z_k, \sum_{i=1}^n F(t_i)\mu(E_i)) \le \varepsilon/2;$$

then

$$\begin{aligned} & \limsup_{n\to\infty} d(z, \sum_{i=1}^n F(t_i)\mu(E_i)) \le \\ \le \ & \limsup_{n\to\infty} \left(\|z - z_k\| + d(z_k, \sum_{i=1}^n F(t_i)\mu(E_i)) \right) \le \varepsilon. \end{aligned}$$

and therefore, by definition, $z \in \Phi(F,[a,b])$.

(b) Moreover, if F is closed and convex valued, $\Phi(F,[a,b])$ is convex too. In fact, since

$$\Phi(F,[a,b]) = \bigcap_{\varepsilon>0} \bigcup_{\Delta} \bigcap_{(E_i,t_i)_i \in \mathcal{P}_\Delta} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^n F(t_i)\mu(E_i), \varepsilon), \qquad (2)$$

if $x, y \in \Phi(F,[a,b])$ then for every $\varepsilon > 0$ there exist Δ_x, Δ_y such that

$$\begin{aligned} x &\in \bigcap_{(E_i,t_i)_i \in \mathcal{P}_{\Delta_x}} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^n F(t_i)\mu(E_i), \varepsilon) \\ y &\in \bigcap_{(E_i,t_i)_i \in \mathcal{P}_{\Delta_y}} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^n F(t_i)\mu(E_i), \varepsilon). \end{aligned}$$

Let $\Delta = \Delta_x \cap \Delta_y$. Then, for every generalized $\mathcal{P}_\Delta$ McShane partition $(E_i,t_i)_i$, we have:

$$x, y \in \bigcap_{(E_i,t_i)_i \in \mathcal{P}_{\Delta}} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^n F(t_i)\mu(E_i), \varepsilon)$$

and so there are two integers m_1, m_2 such that

$$x \in \bigcap_{n=m_1}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon), \quad y \in \bigcap_{n=m_2}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon).$$

If we take $m = \max\{m_1, m_2\}$ then

$$x, y \in \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon)$$

and so, since this last set is convex, for every $a \in [0,1]$,

$$ax + (1-a)y \in \bigcap_{(E_i,t_i)_i \in \mathcal{P}_\Delta} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{U}(\sum_{i=1}^{n} F(t_i)\mu(E_i), \varepsilon).$$

Then the convexity of $\Phi(F, [a,b])$ follows.

(c) If F is integrably bounded, namely there exists $g \in L^1([a,b])$ such that $h(F(t), \{0\}) \leq g(t)$ a.e. , then $\Phi(F, [a,b])$ is bounded. Indeed for every $z \in \Phi(F, [a,b])$ and for every $\varepsilon > 0$ there are a gauge Δ and a point $x \in \sum_{i=1}^{n} F(t_i)\mu(E_i)$ (where $(E_i, t_i)_i$ is a generalized $\mathcal{P}_\Delta$ McShane partition) such that $\|z - x\| \leq \varepsilon$, and hence

$$\|z\| \quad \leq \quad \|z - x\| + \|x\| \leq \varepsilon + \|g\|_1.$$

By the arbitrariness of z, it follows that $\Phi(F, [a,b])$ is bounded.

Observe also that in the definition no separability of X, X' is required. Consider now the classical integral given in the theory of multivalued integration, namely the Aumann integral [1], which is defined by:

$$(A) - \int_a^b F dt = \left\{ (B) - \int_a^b f dt;\ f \in S_F^1 \right\},$$

where S_F^1 is the set of all Bochner integrable selections of F.
We recall also that a multifunction F is measurable if

$$F^-(C) = \{t \in [a,b] : F(t) \cap C \neq \emptyset\}$$

is a Borel set for every closed set $C \subset X$.
We want to compare now the $(*)$- and the (A)-integrals.

Proposition 2 *Let $F : [a,b] \to 2^X \setminus \emptyset$ be a multifunction. Then*

$$(A) - \int_a^b F dt \subset \Phi(F, [a,b]).$$

Proof: Since the proof is easy, we give it only for the sake of simplicity. The inclusion is obvious if the Aumann integral is empty. If it is not, let $z \in (A) \int_a^b F(t)dt$, then there exists a function $f \in S^1_F$ such that $f(t) \in F(t)$ for every $t \in [a,b]$ and $z = \int_a^b f d\mu$. Since f is Bochner integrable it is also McShane integrable and so, for every $\varepsilon > 0$, there exists a gauge Δ such that, for every $\mathcal{P}_\Delta$ generalized McShane partition $(E_i, t_i)_i$, we have

$$\limsup_n d(z, \sum_{i=1}^n F(t_i)\mu(E_i)) \leq \limsup_n \|z - \sum_{i=1}^n f(t_i)\mu(E_i)\| \leq \varepsilon$$

and thus it follows that $z \in \Phi(F, [a,b])$. $\square$.

In order to prove the opposite inclusion we suppose that the multifunction F is also $cwk(X)$-valued, measurable and integrably bounded and that the space X is separable. In this case, we will show that the (A)-integral is non empty. In order to prove this, we recall the following useful results:

Proposition 3 [12, Proposition II.5.20] Let X be a separable Banach space. *If $F : [a,b] \to cwk(X)$ is graph measurable and integrably bounded, then*

$$(A) - \int_a^b F(t)dt \in cwk(X).$$

Proposition 4 [12, Proposition II.5.2] *Let X be a separable Banach space. If $F : [a,b] \to cwk(X)$ is graph measurable and $S^1_F \neq \emptyset$, then for every $x' \in X'$ we have:*

$$s(x', (A) - \int_a^b F(t)dt) = \int_a^b s(x', F(t))\, dt$$

where $s(x', \cdot)$ is the support function defined for any nonempty set $C \subset X$ by $s(x', C) = \sup\{< x', x >: x \in C\}$.

Lemma 1 [4, Lemma III.14] *Let $P = (x'_n)_n$ be a dense sequence in X' for the topology $\tau(X', X)$, and K be a closed, convex, weakly locally compact subset of X which contains no line. Then*

$$K = \cap_n \{x \in X : \ < x'_n, x > \ \leq \ s(x'_n, K)\}.$$

Moreover

Proposition 5 *Let X be a Banach space and $F : [a, b] \to cwk(X)$ be a measurable and integrably bounded multifunction. Then, if we set*

$$L = \int_a^b s(x', F(t))dt$$

for every $x' \in X'$ we have

$$\Phi(F, [a, b]) \subset \{z \in X : \ < x', z > \ \leq \ L\}\,. \tag{3}$$

Proof: Let $z \in \Phi(F, [a, b])$ and suppose that (3) is not true. Then $< x', z > -L = \alpha > 0$.

By definition of $(*)$-integral, there exists a gauge Δ_* such that, for every generalized McShane partition $(E_i, t_i)_i$ subordinate to $\Delta_*(\alpha/6)$, we have:

$$\limsup_n d\left(z, \sum_{i=1}^n F(t_i)\mu(t_i)\right) := r \leq \alpha/6. \tag{4}$$

Let now $\varepsilon > 0$ satisfy $r + \varepsilon < \alpha/3$. Then, in correspondence to ε, there exists an integer $\overline{n}$ such that, for every $n \geq \overline{n}$,

$$d\left(z, \sum_{i=1}^n F(t_i)\mu(E_i)\right) < \alpha/3.$$

Since F has weakly compact values, then, for every $n \geq \overline{n}$, there exists $x_n \in \sum_{i=1}^n F(t_i)\mu(E_i)$ such that $\|z - x_n\| = d(z, \sum_{i=1}^n F(t_i)\mu(E_i))$ and hence

$$\begin{aligned} < x', z > \ &\leq \ | < x', x_n > | + | < x', z - x_n > | \leq \\ &\leq \ s(x', \sum_{i=1}^n F(t_i)\mu(E_i)) + \alpha/3 \leq \\ &\leq \ \sum_{i=1}^n s(x', F(t_i)\mu(E_i)) + \alpha/3. \end{aligned} \tag{5}$$

Moreover we know that $s(x', F)$ is Lebesgue integrable, since it is measurable and dominated by $h(F(t), \{0\})$ and so, by [7, Lemma 1J] already quoted, there exists a gauge Δ_0 such that for every generalized $\mathcal{P}_{\Delta_0}$ McShane partition $\Pi' = (E'_i, t'_i)_i$,

$$\sum_{i=1}^{n} s(x', F(t_i)\mu(E_i)) \leq L + \alpha/3. \tag{6}$$

Therefore, if we consider $\Delta = \Delta_* \cap \Delta_0$ and we take any generalized $\mathcal{P}_\Delta$ McShane partition, inequalities (4), (5) and (6) give us the following contradiction: $< x', z >\leq L + 2\alpha/3 =< x', z > -\alpha/3$. Hence (3) holds.□

Now we are in position to state our comparison result.

Theorem 1 *Suppose that X is a separable Banach space and that there exists a countable family $(x'_n)_n$ in X' which separates points of X. Then*

$$(A)\int_a^b F(t)dt = \Phi(F, [a, b])$$

holds, for any measurable and integrably bounded multifunction $F : [a, b] \to cwk(X)$.

Remark 2 Observe that this theorem extends [14, Theorem 3] in several directions: first of all, we obtain an analogous result in infinite dimensional spaces, rather then a Euclidean space. Moreover here multifunctions with unbounded domains are allowed, and their values are only requested to be convex and weakly-compact. The hypothesis of convexity of the values could not be dropped in our case; indeed in the infinite dimensional case there are examples of non convex Aumann integrals.

Proof of Theorem 1: The inclusion

$$(A) - \int_a^b F dt \subset \Phi(F, [a, b])$$

is contained in Proposition 2 which holds without any assumption on F and X. The other inclusion is proved similarly as in [14, Theorem

3]. Observe that from [4, Lemma III.32] it is possible to construct a countable family P which is dense in E' for $\tau(E', E)$. So, let $P = (x'_n)_n$. By [4, Lemma III.14] quoted above, we know that, for every $t \in [a, b]$,

$$F(t) = \bigcap_n \{z \in X : \ < x'_n, z > \ \leq \ s(x'_n, F(t))\} .$$

Set $L_n = \displaystyle\int_a^b s(x'_n, F(t))dt$. Applying Proposition 5 we have that

$$\Phi(F, [a, b]) \subset \bigcap_n \{z \in X : \ < x'_n, z > \ \leq \ L_n\} .$$

Observe also that, since F is $cwk(X)$-valued, then its (A)-integral belongs to the same hyperspace by Proposition 3 ([12, Proposition 5.20]) and then, using again [4, Lemma III.14], we have

$$(A) - \int_a^b F(t)dt = \bigcap_n \left\{ x \in X :< x'_n, x >\leq \ s(x'_n, (A) - \int_a^b F(t)dt) \right\} .$$

Thus we have proved that

$$\Phi(F, [a, b]) \subset (A) \int_a^b F(t)dt$$

and this concludes the proof of the theorem. □

Theorem 1 can be applied in the comparison of Aumann integral and other known integrals; for the relationship with the Debreu integral see for example [3, 19, 15, 16, 20]. For weakly compact valued multifunctions the result was obtained for totally measurable multifunctions. In general, measurable multifunctions are not totally measurable. Here we give an example of a measurable multifunction not totally measurable for which the Aumann and the $(\star)$-integrals coincide.

Example 1 Let $X = l^2(\mathbb{N}^*)$; for every $A \subset \mathbb{N}^*$ we consider

$$U_A = \{x \in X : \|x\| \leq 1, \text{ and } x_n = 0 \text{ if } \ n \notin A\} = \{1_A x : \|x\| \leq 1\},$$

where $(1_A x)_n = 1_A(n)x_n$. If $A \neq B$ then $h(U_A, U_B) \geq 1$ and so the set $\{U_A, A \subset \mathbb{N}^*\}$ is not separable.

Let $\Omega = [0,1[$ and for every $\omega \in \Omega$ let $0, \omega_1 \cdots \omega_n \cdots$ be its dyadic representation, namely $\omega_1 = 1$ iff $\omega \in [1/2, 1[$, $\omega_2 = 1$ iff $\omega \in [1/4, 1/2[\cup [3/4, 1[$, etc. We set $B_1 = [1/2, 1[, B_2 = [1/4, 1/2[\ \cup \ [3/4, 1[$, etc.

Let $F(\omega) = U_{A(\omega)}$ where $A(\omega) = \{n \in \mathbb{N}^* : \omega_n = 1\}$. F is integrably bounded, takes weakly compact and convex values and its support function $s(y, F(\omega))$ is measurable since it is the limit of simple functions; indeed:

$$s(y, F(\omega)) = \{\sum_{n \in A(\omega)} y_n^2\}^{1/2} = \lim_{n \to \infty} \sum_{p \leq n} s(y, F(\omega)) 1_{B_p(\omega)}$$

and

$$\sum_{p \leq n} s(y, F(\omega)) 1_{B_p(\omega)} = \{\sum_{p \leq n} y_p^2 : \omega_p = 1\}^{1/2}.$$

From [12, Proposition II.2.39] F is measurable, but for every μ-null set N, the set $\Omega \setminus N$ is not countable and so $F(\Omega \setminus N)$ is not separable in the h-metric topology. Then immediately it follows that F cannot be a member of the closure of simple multifunctions with weakly compact and convex values in the L^1-metric associated with h and so F is not a Bochner integrable multifunction. Moreover, by Theorem 1,

$$\Phi(F, [0,1[) = (A) - \int_0^1 F(\omega) d\mu(\omega).$$

4. The McShane multivalued integral

If we consider directly the hyperspace $(cwk(X), h)$ we can define the McShane multivalued integral in the following way:

Definition 3 We say that $F : [a,b] \to cwk(X)$ is *McShane integrable* if there exists $J \in cwk(X)$ such that for every $\varepsilon > 0$ there exists a gauge

Δ such that

$$\limsup_n h(J, \sum_{i=1}^n F(t_i)\mu(E_i)) \leq \varepsilon$$

for every generalized $\mathcal{P}_\Delta$ McShane partition $\Pi = (E_i, t_i)_i$. In this case, we set

$$J := \int_a^b F(t)dt.$$

Thanks to the Rådstrom embedding theorem [18], this definition is well-posed, and we will show the following:

Theorem 2 *If $F : [a, b] \to cwk(X)$ is McShane integrable, then the $(\star)$-integral and the McShane integral coincide, namely $J = \Phi(F, [a, b])$.*

Proof: The inclusion $J \subset \Phi(F, [a, b])$ is obvious; indeed if $z \in J$, then for every $\varepsilon > 0$ there exists a gauge Δ such that for each generalized $\mathcal{P}_\Delta$ McShane partition $(E_i, t_i)_{i\in\mathbb{N}}$ we get:

$$\limsup_n d(z, \sum_{i=1}^n F(t_i)\mu(E_i)) \quad \leq \quad \limsup_n h(J, \sum_{i=1}^n F(t_i)\mu(E_i)) \leq \varepsilon.$$

Conversely, let now $z \in \Phi(F, [a, b])$. Then, for every $\varepsilon > 0$, there exists a gauge Δ such that, for every generalized $\mathcal{P}_\Delta$ McShane partition $(E_i, t_i)_{i\in\mathbb{N}}$, we have

$$\limsup_n d(z, \sum_{i=1}^n F(t_i)\mu(E_i)) \leq \varepsilon/2.$$

On the other hand, by the definition of the McShane integral, there exists a gauge Δ_1 such that for every generalized $\mathcal{P}_{\Delta_1}$ McShane partition $(E'_i, t'_i)_i$, we get

$$\limsup_n h(J, \sum_{i=1}^n F(t'_i)\mu(E'_i)) \leq \varepsilon/2.$$

So, if we take $\widetilde{\Delta} = \Delta \cap \Delta_1$, then, for every generalized $\mathcal{P}_{\widetilde{\Delta}}$ McShane partition $\Pi = (E_i, t_i)_i$ and for every $x \in \sum_{i=1}^n F(t_i)\mu(E_i)$ we have

$$\begin{aligned} d(z, J) &= \inf_{y \in J} \|z - y\| \leq \inf_{y \in J} (\|z - x\| + \|x - y\|) = \\ &= \|z - x\| + d(x, J) \leq \|z - x\| + h(\sum_{i=1}^n F(t_i)\mu(E_i), J). \end{aligned}$$

So we have

$$d(z, J) \leq \limsup_n \left(d(z, \sum_{i=1}^n F(t_i)\mu(E_i)) + h(\sum_{i=1}^n F(t_i)\mu(E_i), J) \right) \leq \varepsilon$$

for every generalized $\mathcal{P}_{\widetilde{\Delta}}$ McShane partition $(E_i, t_i)_i$. Since ε is arbitrary and $\Phi(F, [a, b])$ is closed, the last inclusion follows. □

Observe that if X is separable and $F : \Omega \to ck(X)$ is a measurable multifunction with unbounded range, then Debreu integrability implies McShane's one. In this case we can embed $(ck(X), h)$ in a suitable separable Banach space Y and, if we consider F as a Y-valued measurable function, McShane integrability coincides with the Pettis' one. If the range of F is bounded and $\mu(\Omega) < \infty$, the two concept of integral coincide; see for example [6, Section 2K].
If we consider a multifunction with weakly compact and convex values we need total measurability of F since $(cwk(X), h)$ is not separable in general. In this case we have:

Corollary 1 *If μ is finite, X is a separable Banach space and $F : \Omega \to cwk(X)$ is a Debreu integrable multifunction, then F is McShane integrable and its integral coincides with the Aumann integral of F.*

Proof: Thanks to the result of Byrne [3] the Debreu integral and the Aumann integral coincide; thank to [7, Theorem 1K] the Debreu and the McShane integrals coincide too. □

An example of an integrably bounded and McShane integrable multifunction which is not Debreu integrable can be obtained using [6, Example 3F] and taking $F(t) = \{f(t)\}$, where f takes values in the non separable Banach space $L^\infty([0,1])$.

Theorem 2 implies that, when F is $cwk(X)$-valued and McShane integrable, $\Phi(F, [a,b])$ is convex and weakly compact and so in this case we obtain [14, Proposition 1] as a corollary of Theorem 2.
Theorem 2 is also important from another point of view: indeed, thanks to the Rådstrom embedding theorem, all the fundamental results concerning the McShane integral, which are given in [7], are still valid for $cwk(X)$-valued multifunctions. So it is enough to consider the space $cwk(X)$ as a Banach space, which plays the role of the Banach space X in the previous section. So $\Phi(F, [a,b])$ satisfies the main fundamental properties of the functionals defined by means of integrals, like for example additivity and absolute continuity.

Remark 3 Though in this paper Ω is always assumed to be an interval in the real line (possibly unbounded), we observe that all the results here obtained hold as well whenever Ω is any non empty σ-finite quasi Radon outer regular measure space.

Acknowledgement

The authors would like to thank Prof. Jan Andres for the interesting discussions on the argument and Prof. Domenico Candeloro for his kind support, his helpful suggestions and comments during this work.

This work was partially supported by the G.N.A.M.P.A. of C.N.R.

References

[1] R. J. AUMANN, Integrals of set-valued functions, *J. Math. Anal. Appl.* **12** 1-12, (1965).

[2] D. BARCENAS - W. URBINA, Measurable multifunctions in non-separable Banach spaces, *SIAM J. Math. Anal.* **28,** N. 5 (1997), 1212-1226.

[3] C. L. BYRNE Remarks on the Set-Valued Integrals of Debreu and Aumann,*J. Math. Anal. Appl.* **62**, 243-246, (1978).

[4] C. CASTAING - M. VALADIER, Convex Analysis and Measurable Multifunctions, *Lecture Notes in Math.* **580***, Springer-Verlag* (1977).

[5] J. DIESTEL - J. J. UHL, Vector measures, *Mathematical Surveys - Number 15 Amer. Math. Soc. Providence,* Rhode Island (1977).

[6] D. H. FREMLIN - J. MENDOZA, On the integration of vector-valued functions, *Ill. J. Math.* **38** (1994), 127-147.

[7] D. H. FREMLIN, The generalized McShane integral, *Ill. J. Math.* **39** (1995), 39-67.

[8] R. GORDON The McShane Integral of Banach valued functions, *Illinois J. of Math.,* **34** 3, (1990) 557-567.

[9] C. HESS, Sur la measurabilité des multifunctions à valeurs faiblement compactes sans droites, *C. R. Acad. Sci. Paris,* **305** Serie I, (1987) 631-634.

[10] C. HESS, Measurability and Integrability of the Weak Upper Limit of a Sequence of Multifunctions, *J. Math. Anal. Appl.,* **153**, (1990) 226-249.

[11] C.J. HIMMELBERG Measurable relations, *Fund. Math.,* **87** (1975) 53-72.

[12] S. HU - N. S. PAPAGEORGIOU, Handbook of Multivalued Analysis - Volume I: Theory, in *Mathematics and Its Applications,* **419***, Kluwer Academic Publisher,* Dordrecht (1997).

[13] K. KURATOWSKI - C. RYLL-NARDZEWSKI A general theorem on selectors, *Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys.,* **13** (1965), 397-403.

[14] J. J. JARNIK - J. KURZWEIL, Integral of multivalued mappings and its connection with differential relations, *Časopis pro pěstování matematiky* **108** (1983), 8-28.

[15] A. MARTELLOTTI - A. R. SAMBUCINI, On the comparison between Aumann and Bochner integrals, *J. Math. Anal. Appl.* **260** N. 1 (2001), 6-17.

[16] A. MARTELLOTTI - A. R. SAMBUCINI The finitely additive integral of multifunctions with closed and convex values, *Zeitschrift fur Analysis ihre Anwendungen,* **21** N. 4 (2002), 851-864.

[17] T. NEUBRUNN - B. RIEČAN, Integral, Measure, and Ordering, *Mathematics and Its Applications,* **411***, Kluwer Academic Publisher,* Dordrecht (1997).

[18] H. RÅDSTROM, An Embedding Theorem for Spaces of Convex Sets, *Proc. Amer. Math. Soc.* **3** (1952), 165-169.

[19] A. R. SAMBUCINI, Remarks on set valued integrals of multifunctions with non empty, bounded, closed and convex values, *Commentationes Math.* **39**, (1999), 153-165.

[20] A. R. SAMBUCINI, A survey on multivalued integration, *Atti Mat. Fis. Univ. Modena,* **50** (2002), 53-63.

[21] M. TALAGRAND, Pettis integral and measure theory, *Mem. Amer. Math. Soc.,* **307***,* (1984).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.4,327-367,2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

A General Framework for Term Structure Models Driven by Lévy Processes

Jorge Hernández
Department of Statistics and Applied Probability
University of California
Santa Barbara, CA 93106
jorge@pstat.ucsb.edu

Abstract

We describe a framework in which to generalize the Heath, Jarrow and Morton model for the term structure of interest rates. We represent the model in terms of the triplet of characteristics of the underlying semimartingales. We state and prove the necessary and sufficient conditions for absence of arbitrage in terms of the characteristics of the price process. The methodology is then extended to find sufficient conditions for absence of arbitrage in the defaultable case.

Keywords semimartingales, finance, term structure of interest rates

2000 AMS Subject Classification 60G51, 91B28, 91B70

1 Introduction

Our goal is to present a general framework in which to construct bond market models. We introduce the class of semimartingales and a stochastic calculus associated with it. We borrow this material from Shiryaev and Jacod [11] and we adopt their notation. We outline the bond market setting presented in Björk, Kabanov and Runggaldier [2] and give an alternative proof of the main result regarding the absence of arbitrage, which extends to the case where the Levy measure of the driving process is infinite.

Shirakawa [10] introduced an extension of the HJM approach for term structure models that allows for jumps in the forward rate dynamics. The model is driven by a Poisson process with constant intensities, in addition to a standard Brownian Motion. Jarrow, Madan [6] define a multivariate point process in terms of a sequence of stopping times at which jumps take place. The associated counting process is used as one of the driving terms. Björk, et al. [2] extend the HJM model to infinite jump space by introducing a random measure with finite compensator.

Babbs and Webber [1] and El-Jahel, Lindberg and Perraudin [3] use an alternative approach that takes into account monetary policy. The short rate is assumed to be established periodically by the authorities and hence is modelled by a pure jump process. In the latter, the intensities follow a squared Ornstein-Uhlenbeck process. In Babbs and Webber [1] the intensities depend on the short rate itself and in a Markov process which represents the state of the economy.

The class of semimartingales is particularly useful for our purpose since it includes a large variety of processes. In connection with the model in Björk, et al. [2] we will work with a special case, namely diffusion processes. Further, the class of semimartingales is invariant with respect many transformations. Ito's formula, for example, implies invariance with respect to composition with C^2 functions. The results involving arbitrage depend heavily on invariance after an absolutely continuous change of measure.

We begin the exposition by introducing some definitions and the notation that will be used throughout.

A *stochastic basis* is a probability space $(\Omega, \mathscr{F}, \mathbf{F}, P)$ equipped with a *filtration* $\mathbf{F}= (\mathscr{F}_t)_{t\in\mathbb{R}_+}$; here *filtration* means an increasing and right-continuous family of sub-σ-fields of $\mathscr{F}$. By convention, we set $\mathscr{F}_\infty = \mathscr{F}$ and $\mathscr{F}_{\infty-} = \bigvee_{s\in\mathbb{R}_+}\mathscr{F}_s$.

The stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is called *complete*, or equivalently is said to *satisfy the usual conditions* if the σ-field $\mathscr{F}$ is P-complete and if every $\mathscr{F}_t$ contains all the P-null sets of $\mathscr{F}$.

A *random set* is a subset of $\Omega \times \mathbb{R}_+$.

A *process*(or, an *E-valued process* is a family $X = (X_t)_{t\in\mathbb{R}_+}$ of mappings from Ω into some set E. Unless otherwise stated, E will be $\mathbb{R}^d$ for some $d \in \mathbb{N}^*$.

A process X is called *càd* (resp.*càg*, resp.*càdlàg*) if all its paths are right-continuous (resp. are left- continuous, resp. are right-continuous and admit left-hand limits). When X is càdlàg we define two other processes $X_- = (X_{t-})_{t\in\mathbb{R}_+}$ and $\Delta X = (\Delta X_t)_{t\in\mathbb{R}_+}$ by

$$
\begin{aligned}
&X_{0-} = X_0, \quad X_{t-} = \lim_{s\to t, s<t} X_s \text{ for } t > 0 \\
&\Delta X_t = X_t - X_{t-}.
\end{aligned}
$$

If X is a process and if T is a mapping: $\Omega \to \overline{\mathbb{R}}_+$, we define the *process stopped at time* T, denoted by X^T, by $X_t^T = X_{T\wedge t}$.

A random set A is called *evanescent* if the the set $\{\omega : \exists t \in \mathbb{R}_+ \text{ with } (\omega, t) \in A\}$ is P-null. Two processes X and Y are called *indistinguishable* if the random set $\{X \neq Y\} = \{(\omega, t) : \ X_t(\omega) \neq Y_t(\omega)\}$ is evanescent.

In what follows, let $(\Omega, \mathscr{F}, \mathbf{F}, P)$ be a stochastic basis.

1.1 σ-Fields, Random Times

A process X is *adapted* to the filtration $\mathbf{F}$ if X_t is $\mathscr{F}_t$-measurable for every $t \in \mathbb{R}_+$.

A *stopping time* is a mapping $T : \Omega \to \overline{\mathbb{R}}_+$ such that $\{T \leq t\} \in \mathscr{F}_t$ for all $t \in \mathbb{R}_+$.

If T is a stopping time, we denote by $\mathscr{F}_T$ the collection of all sets $A \in \mathscr{F}$ such that $A \cap \{T \leq t\} \in \mathscr{F}_t$ for all $t \in \mathbb{R}_+$.

If T is a stopping time, we denote by $\mathscr{F}_{T-}$ the σ-field generated by $\mathscr{F}_0$ and all the sets of the form $A \cap \{t < T\}$, where $t \in \mathbb{R}_+$ and $A \in \mathscr{F}_t$.

The *optional σ-field* is the σ-field $\mathscr{O}$ on $\Omega \times \mathbb{R}_+$ that is generated by all *càdlàg* adapted processes (considered as mappings on $\Omega \times \mathbb{R}_+$). A process or random set that is $\mathscr{O}$-measurable is called *optional*.

Proposition 1.1 *Let X be an optional process. When considered as a mapping on $\Omega \times \mathbb{R}_+$, it is $\mathscr{F} \otimes \mathscr{R}_+$- measurable. Moreover, if T is a stopping time, then*

a) $X_T 1_{\{T<\infty\}}$ is $\mathscr{F}_T$-adapted (hence, X is adapted).

b) the stopped process X^T is also optional.

Let S, T be two stopping times. We define the *stochastic intervals* to be the following random sets:

$$
\begin{aligned}
[\![S,T]\!] &= \{(\omega,t)\colon t \in \mathbb{R}_+, S(\omega) \leq t \leq T(\omega)\} \\
[\![S,T[\![&= \{(\omega,t)\colon t \in \mathbb{R}_+, S(\omega) \leq t < T(\omega)\} \\
]\!]S,T]\!] &= \{(\omega,t)\colon t \in \mathbb{R}_+, S(\omega) < t \leq T(\omega)\} \\
]\!]S,T[\![&= \{(\omega,t)\colon t \in \mathbb{R}_+, S(\omega) < t < T(\omega)\}
\end{aligned}
$$

We will denote the stochastic interval $[\![T,T]\!]$ by $[\![T]\!]$ and will call it the *graph of the stopping time T*.

Proposition 1.2 *If S, T are two stopping times and if Y is an $\mathscr{F}_S$-measurable random variable, the processes $Y1_{[\![S,T]\!]}$, $Y1_{[\![S,T[\![}$, $Y1_{]\!]S,T]\!]}$ and $Y1_{]\!]S,T[\![}$ are optional.*

Let X be an adapted process. For each $n \in \mathbb{N}^*$ define a new process X^n by

$$X^n = \sum_{k \in \mathbb{N}} X_{k/2^n} 1_{[\![k/2^n,(k+1)/2^n[\![}.$$

Note that if X is càg then the sequence (X^n) converges pointwise to X. Hence Proposition 1.2 yields the following result:

Proposition 1.3 *Every process X that is càg and adapted is optional.*

Corollary 1.1 *If X is a càdlàg adapted process, the processes X_- and ΔX are optional.*

A random set A is called *thin* if it is of the form $A = \bigcup_n [\![T_n]\!]$, where (T_n) is a sequence of stopping times. If moreover the sequence (T_n) satisfies $[\![T_n]\!] \cap [\![T_m]\!] = \emptyset$ for all $n \neq m$, it is called an *exhausting sequence* for A.

Lemma 1.1 *Any thin random set admits an exhausting sequence of stopping times.*

Proposition 1.4 *If X is a càdlàg adapted process, the random set $\{\Delta X \neq 0\}$ is thin. An exhausting sequence for this set is called a sequence that* exhausts the jumps of X.

Observe that if X is càdlàg then for each $\omega \in \Omega$ the set $\{t \in \mathbb{R}_+ : \Delta X_t(\omega) \neq 0\}$ is countable. Proposition 1.4 and Lemma 1.1 will be applied as follows: Suppose $(T_n)_n$ is a sequence of stopping times that exhausts the thin set $\{\Delta X \neq 0\}$. Then if $(\omega, t) \in \{\Delta X \neq 0\}$ there is a unique n such that $T_n(\omega) = t$. This will be useful in analyzing the "jumps" of a process describing the behavior of asset prices.

The *predictable σ-field* is the σ-field $\mathscr{P}$ on $\Omega \times \mathbb{R}_+$ that is generated by all *càg* adapted processes (considered as mappings on $\Omega \times \mathbb{R}_+$). A process or random set that is $\mathscr{P}$- measurable is called *predictable*. Note that Proposition 1.3 implies that $\mathscr{P} \subset \mathscr{O}$.

Theorem 1.1 *The predictable σ-field is also generated by any one of the following collections of sets:*

(i) $A \times \{0\}$ *where* $A \in \mathscr{F}_0$, *and* $[\![0, T]\!]$ *where* T *is any stopping time;*

(ii) $A \times \{0\}$ *where* $A \in \mathscr{F}_0$, *and* $A \times (s, t]$ *where* $s < t, A \in \mathscr{F}_s$.

Proposition 1.5 *If* X *is a predictable process and if* T *is a stopping time,*

a) $X_T 1_{\{T<\infty\}}$ *is* $\mathscr{F}_{T-}$*-measurable,*

b) the stopped process X^T *is also predictable.*

Proposition 1.6 *If* S, T *are two stopping times and if* Y *is an* $\mathscr{F}_S$*-measurable random variable, then the process* $Y 1_{]\!]S,T]\!]}$ *is predictable.*
process is adapted and càg)

Proposition 1.7 *If* X *is a càdlàg adapted process, then* X_- *is a predictable process. If moreover* X *is predictable, then* ΔX *is predictable.*
immediate from the definition, since X_- *is adapted and càg)*

A *predictable time* is a mapping $T\colon \Omega \to \overline{\mathbb{R}}_+$ such that the stochastic interval $[\![0, T[\![$ is predictable.

Theorem 1.2 *a) Let* X *be an* $\overline{\mathbb{R}}$*-valued and* $\mathscr{F} \otimes \mathscr{R}_+$*-measurable process. There exists a* $(-\infty, \infty]$*-valued process, called the* predictable projection *of* X *and denoted by* ${}^pX$, *that is determined uniquely up to an evanescent set by the following two conditions:*

(i) it is predictable,

(ii) $({}^pX)_T = E(X_T|\mathscr{F}_{T-})$ *on* $\{T < \infty\}$ *for all predictable times* T.

b) Moreover, if T *is any stopping time, then*

$$
{}^p(X^T) = ({}^pX)1_{[\![0,T]\!]} + X_T 1_{]\!]T,\infty[\![}.
$$

c) Moreover, if ${}^pX$ *is finite-valued and if* X' *is a* $(-\infty, \infty]$*- valued predictable process, then* ${}^p(XX') = X'{}^p(X)$.

Note that if M is a local martingale then ${}^pM = M_-$ and ${}^p(\Delta M) = 0$. Further, if $X = X_0 + M + A$ is a special semimartingale so that A is predictable, then ${}^pX - X_- = \Delta A$. In particular, ${}^pX = X_-$ if and only if $\Delta A = 0$.

If $\mathscr{C}$ is a class of processes, we denote by $\mathscr{C}_{\mathrm{loc}}$ its *localized class*, defined as such: a process X belongs to $\mathscr{C}_{\mathrm{loc}}$ if and only if there exists an increasing sequence (T_n) of stopping times (depending on X) such that $\lim_n T_n = \infty$ a.s. and that each stopped process X^{T_n} belongs to $\mathscr{C}$. The sequence (T_n) is called a *localizing sequence* for X (relative to $\mathscr{C}$).

1.2 Martingales

In this subsection we state some standard results concerning martingales.

A *martingale*(resp. *submartingale*; resp. *supermartingale*) is an adapted process X on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, whose P-almost all paths are *càdlàg*, such that every X_t is integrable, and that for $s \leq t$:

$$X_s = E(X_t|\mathscr{F}_s) \quad (\text{resp. } X_s \leq E(X_t|\mathscr{F}_s);\ \text{resp. } X_s \geq E(X_t|\mathscr{F}_s)).$$

We say that a process X admits a *terminal variable* X_∞ if X_t converges a.s. to a limit X_∞ as $t \uparrow \infty$; in which case the variable X_T is (a.s.) well defined for any stopping time T, with $X_T = X_\infty$ on $\{T < \infty\}$.

Theorem 1.3 *Let X be a supermartingale such that there exists an integrable random variable Y with $X_t = E(Y|\mathscr{F}_t)$ for all $t \in \mathbb{R}_+$. Then*

a) (Doob's limit Theorem) X_t converges a.s. to a finite limit X_∞.

b) (Doob's stopping Theorem) If S, T are stopping times, the random variables X_S and X_T are integrable, and $X_S \geq E(X_T|\mathscr{F}_S)$ on the set $\{S \leq T\}$. In particular, X^T is again a supermartingale.

We denote by $\mathscr{M}$ the class of all *uniformly integrable martingales*; and denote by $\mathscr{H}^2$ the class of all *square-integrable martingales*, that is, of all martingales X such that $\sup_{t\in\mathbb{R}_+} E(X_t^2) < \infty$.

Theorem 1.4 *a) If X is a uniformly integrable martingale, then X_t converges a.s. and in L^1 to a terminal variable X_∞, and $X_T = E(X_\infty|\mathscr{F}_T)$ for all stopping times T. Moreover, X is square-integrable if and only if X_∞ is square-integrable, in which case the convergence $X_t \to X_\infty$ also takes place in L^2.*

b) If Y is an integrable random variable, there exists a uniformly integrable martingale X, and only one up to an evanescent set, such that $X_t = E(Y|\mathscr{F}_t)$ for all $t \in \mathbb{R}_+$; moreover, $X_\infty = E(Y|\mathscr{F}_{\infty-})$.

Theorem 1.5 (Doob's inequality) *If X is a square-integrable martingale,*

$$E\left(\sup_{t\in\mathbb{R}_+} X_t^2\right) \leq 4 \sup_{t\in\mathbb{R}_+} E(X_t^2) = 4E(X_\infty^2).$$

The following result gives a useful characterization of uniformly integrable martingales.

Lemma 1.2 *Let X be an adapted càdlàg process, with a terminal random variable X_∞. Then X is a uniformly integrable martingale if and only if for each stopping time T, the variable X_T is integrable and satisfies $E(X_T) = E(X_0)$.*

A *local martingale*(resp. a *locally square-integrable martingale*) is a process that belongs to the localized class $\mathscr{M}_{\mathrm{loc}}$ (resp. $\mathscr{H}^2_{\mathrm{loc}}$).

1.3 Increasing Processes, Decomposition of Local Martingales

We now present some further concepts that will be needed in order to introduce the class of semimartingales. First, the definition of the integral of an optional process with respect to an adapted process of finite variation is presented. The following subsection will describe the construction of a more general "stochastic integral" with respect to a semimartingale. Then we state some important results concerning the decomposition of a local martingale. Subsequent sections will state results which characterize the structure of the individual components.

We denote by $\mathscr{V}^+$ (resp.$\mathscr{V}$) the set of all real-valued processes A that are *càdlàg*, adapted, $A_0 = 0$, and whose paths are non-decreasing (resp. have finite variation over each finite interval $[0, t]$).

Let $A \in \mathscr{V}$. We denote by $\mathrm{Var}(A)$ the *variation process of* A, defined by

$$\mathrm{Var}(A)_t(\omega) = \lim_n \sum_{1 \le k \le n} |A_{tk/n}(\omega) - A_{t(k-1)/n}(\omega)|.$$

Suppose $A \in \mathscr{V}$. For each $\omega \in \Omega$, the path $t \to A_t(\omega)$ is the distribution function of a signed measure on $\mathbb{R}^+$ (denoted $dA_t(\omega)$) that is finite on each finite interval $[0, t]$, and is finite on $\mathbb{R}^+$ if and only if $\mathrm{Var}(A)_\infty(\omega) < \infty$.

If $A \in \mathscr{V}$ and H is an optional process, we can define the *integral process* $H \cdot A$ by

$$H \cdot A_t(\omega) = \int_0^t H_s(\omega) dA_s(\omega) \tag{1}$$

if $\int_0^t |H_s(\omega)|\, d[\mathrm{Var}(A)]_s(\omega) < \infty$.

We denote by $\mathscr{A}^+$ the set of all $A \in \mathscr{V}^+$ that are integrable: $\mathrm{E}(A_\infty) < \infty$; and denote by $\mathscr{A}$ the set of all $A \in \mathscr{V}$ that have *integrable variation*: $\mathrm{E}(\mathrm{Var}(A)_\infty) < \infty$.

Theorem 1.6 *Let $A \in \mathscr{A}_{loc}$. There exists a process, called the* compensator *of A and denoted by A^p, which is unique up to an evanescent set, and which is characterized by being a predictable process of $\mathscr{A}_{loc}$ such that $A - A^p$ is a local martingale. Moreover, for each predictable process H such that $H \cdot A \in \mathscr{A}_{loc}$ then $H \cdot A^p \in \mathscr{A}_{loc}$ and $H \cdot A^p = (H \cdot A)^p$. In particular, $H \cdot A - H \cdot A^p$ is a local martingale.*

The following are some easy properties of the compensator.

1. If $A \in \mathscr{A}_{\text{loc}}$ is predictable, then $A^p = A$.
2. If $A \in \mathscr{A}_{\text{loc}}$, then ${}^p(\Delta A) = \Delta(A^p)$.
3. If $A \in \mathscr{A}_{\text{loc}}$, then A is a local martingale if and only if $A = 0$.

Theorem 1.7 *To each pair (M, N) of locally square-integrable martingales one associates a predictable process $\langle M, N\rangle \in \mathscr{V}$, unique up to an evanescent set, such that $MN - \langle M, N\rangle$ is a local martingale. Moreover,*

$$\langle M, N\rangle = \frac{1}{4}\left(\langle M+N, M+N\rangle - \langle M-N, M-N\rangle\right),$$

and if $M, N \in \mathscr{H}^2$ then $\langle M, N\rangle \in \mathscr{A}$ and $MN - \langle M, N\rangle \in \mathscr{M}$. Furthermore, $\langle M, M\rangle$ is nondecreasing.
only if M is quasi-left-continuous.

The process $\langle M, N\rangle$ is called the *predictable quadratic covariation* (also the *angle bracket*) of the pair (M, N). Note that $\langle M, N\rangle = \langle M - M_0, N - N_0\rangle$.

There is a bijective correspondence between the elements M of $\mathscr{H}^2$ and their terminal variables M_∞. If we define the inner product and the norm on $\mathscr{H}^2$ by

$$(M, N)_{H^2} = \mathrm{E} = (M_\infty N_\infty), \quad \|M\|_{H^2} = \|M_\infty\|_{L^2} \text{ where } M, N \in \mathscr{H}^2,$$

then $\mathscr{H}^2$ is a Hilbert space. To see this, let (M^n) be a Cauchy sequence for $\|\cdot\|_{H^2}$. Then the sequence (M^n_∞) is Cauchy in $L^2(\Omega, \mathscr{F}_{\infty-}, P)$; let M_∞ be its L^2 limit. If M is the (unique) martingale with terminal variable M_∞, then it belongs to $\mathscr{H}^2$ and $\|M^n - M\|_{H^2} \to 0$.

Note that the previous theorem implies $(M, N)_{H^2} = \mathrm{E}(\langle M, N\rangle_\infty) + \mathrm{E}(M_0 N_0)$.

Corollary 1.2 *The set of all continuous elements of $\mathscr{H}^2$ is a closed subspace of the Hilbert space $\mathscr{H}^2$.*

Example (The Wiener process).

a) A *Wiener process* on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ (or, relative to $\mathbf{F}$) is a continuous adapted process W such that $W_0 = 0$ and

(i) $\mathrm{E}(W_t^2) < \infty$ for each $t \in \mathbb{R}_+$, and $\mathrm{E}(W_t) = 0$ for each $t \in \mathbb{R}_+$;

(ii) $W_t - W_s$ is independent of the σ-field $\mathscr{F}_s$ for all $0 \leq s \leq t$.

b) The function $\sigma^2(t) = \mathrm{E}(W_t^2)$ is called the *variance function* of W. If $\sigma^2(t) = t$, we say that W is a *standard Wiener process*.

Proposition 1.8 *A Wiener process W is a continuous martingale, and its angle bracket $\langle W, W \rangle$ is $\langle W, W \rangle_t(\omega) = \sigma^2(t)$.*

We now turn to decompositions of a local martingale.

Two local martingales M and N are called *orthogonal* if their product MN is a local martingale. A local martingale X is called a *purely discontinuous local martingale* if $X_0 = 0$ and if it is orthogonal to all continuous local martingales.

Lemma 1.3 *A local martingale that belongs to $\mathscr{V}$ is purely discontinuous.*

Proposition 1.9 *Let $M, N \in \mathscr{H}^2$. The following are equivalent:*

a) M and N are orthogonal.

b) $\langle M, N \rangle = 0$.

c) For all stopping times T, M^T and $N - N_0$ are orthogonal in the Hilbert space $\mathscr{H}^2$.

Corollary 1.3 *The set $\mathscr{H}^{2,\mathrm{d}}$ of all purely discontinuous martingales in $\mathscr{H}^2$ is the orthogonal subspace, in the Hilbert space $\mathscr{H}^2$, of the set $\mathscr{H}^{2,\mathrm{c}}$ of all continuous elements of $\mathscr{H}^2$.*

Proposition 1.10 *Let $a > 0$. Any local martingale M admits a (non-unique) decomposition $M = M_0 + M' + M''$, where M' and M'' are local martingales with $M'_0 = M''_0 = 0$, M' has finite variation, and $|\Delta M''| \leq a$ (hence $M'' \in \mathscr{H}^2_{loc}$).*

Since the local martinagale M' in Proposition 1.10 is of finite variation then it is purely discontinuous by Lemma 1.3. Note also that by Corollary 1.3 $M'' \in \mathscr{H}_{\mathrm{loc}}$ is the sum of purely discontinuous and continuous components. We thus obtain the following decomposition:

Theorem 1.8 *Any local martingale M admits a unique (up to indistinguishability) decomposition*

$$M = M_0 + M^c + M^d,$$

where $M^c = 0$, $M^d = 0$, M^c is a continuous local martingale, and M^d is a purely discontinuous local martingale.

M^c is called the *continuous part* of M, and M^d its *purely discontinuous part*. We denote by $\mathscr{L}$ the set of all local martingales M such that $M_0 = 0$.

1.4 Semimartingales, Stochastic Integrals

Here we introduce the class of semimartingales, which is our basic modeling tool. Hence we are interested in developing a calculus with respect to semimartingales. In this section we give meaning to the notion of a stochastic integral and state some of its properties.

A *semimartingale* is a process X of the form

$$X = X_0 + M + A, \quad M \in \mathscr{L}, \quad A \in \mathscr{V}, \tag{2}$$

where X_0 is finite-valued and $\mathscr{F}_0$-measurable. We denote by $\mathscr{S}$ the space of all semimartingales. A *special semimartingale* is a semimartingale X which admits a decomposition (2) where the process A is predictable. In this case, A is unique (up to an evanescent set). We denote by $\mathscr{S}_p$ the set of all special semimartingales.

If $X \in \mathscr{S}_p$, the unique decomposition $X = X_0 + M + A$ such that $M \in \mathscr{L}$ and that A is a predictable element of $\mathscr{V}$ is called the *canonical decomposition of X*.

Proposition 1.11 *Let X be a semimartingale. The following are equivalent:*

(i) X is a special semimartingale;

(ii) there exists a decomposition $X = X_0 + M + A$ where $A \in \mathscr{A}_{loc}$;

(iii) all decompositions $X = X_0 + M + A$ satisfy $A \in \mathscr{A}_{loc}$;

(iv) the process $Y_t = \sup_{s \leq t} |X_s - X_0|$ belongs to $\mathscr{A}^+_{loc}$.

Lemma 1.4 *If a semimartingale X satisfies $|\Delta X| \leq a$, then it is special and its canonical decomposition $X = X_0 + M + A$ satisfies $|\Delta A| \leq a$ and $|\Delta M| \leq 2a$. In particular, if X is continuous then M and A are continuous.*

Proposition 1.12 *Let X be a semimartingale. There is a unique (up to indistinguishability) continuous local martingale X^c with $X_0^c = 0$, such that any decomposition of type (2) meets $M^c = X^c$ (up to indistinguishability). X^c is called the* continuous martingale part *of X.*

We will now construct the integral process $H \cdot X$ where H is a locally bounded predictable process and X is a semimartingale. Denote by $\mathscr{E}$ the set of all processes of the following forms:

$$H = Y1_{[\![0]\!]},\ Y \text{ is bounded } \mathscr{F}_0\text{-measurable, and}$$
$$H = Y1_{]\!]r,s]\!]},\ r < s,\ Y \text{ is bounded } \mathscr{F}_r\text{-measurable.}$$

For $H \in \mathscr{E}$ and X a semimartingale, the integral process $H \cdot X_t$ is defined by

$$H \cdot X_t = \begin{cases} 0 & \text{if } H = Y1_{[\![0]\!]}, \\ Y(X_{s\wedge t} - X_{r\wedge t}) & \text{if } H = Y1_{]\!]r,s]\!]}. \end{cases} \tag{3}$$

Theorem 1.9 *Let X be a semimartingale. The map $H \to H \cdot X$ defined on $\mathscr{E}$ by (3) has an extension, still denoted by $H \to H \cdot X$ to the space of all locally bounded predictable processes H, with the following properties:*

i) $H \cdot X$ is a càdlàg adapted process;

ii) $H \to H\cdot X$ is linear, up to evanescence (i.e. $(aH+K)\cdot X$ and $aH\cdot X+K\cdot X$ are indistinguishable);

iii) if a sequence (H^n) of predictable processes converges pointwise to a limit H, and if $|H^n| \leq K$ where K is a locally bounded predictable process, then $H^n \cdot X_t \to H \cdot X_t$ in measure for all $t \in \mathbb{R}_+$.

Moreover, this extension is unique, up to evanescence (i.e. if $H \to \alpha(H)$ is another extension with the same properties, then $\alpha(H)$ and $H \cdot X$ are indistinguishable), and in iii) above $H^n \cdot X$ converges to $H \cdot X$ in measure, uniformly on finite intervals: $\sup_{s\leq t}|H^n \cdot X_s - H \cdot X_s| \xrightarrow{P} 0$.

We now state some elementary properties of stochastic integrals. Here X is a semimartingale, and H, K are locally bounded predictable process. All statements are up to evanescence.

(P1) $X \to H \cdot X$ is linear.

(P2) (a) $H \cdot X$ is a semimartingale;

(b) if X is a local martingale then so is $H \cdot X$;

(c) if $X \in \mathscr{V}$ then $H \cdot X \in \mathscr{V}$ and $H \cdot X$ coincides with the process defined in (1) (Stieltjes integral process).

(P3) $(H \cdot X)_0 = 0$ and $H \cdot X = H \cdot (X - X_0)$.

(P4) $\Delta(H \cdot X) = H\Delta X$.

(P5) $X^T = X_0 + 1_{[\![0,T]\!]} \cdot X$ and $(H \cdot X)^T = (H1_{[\![0,T]\!]}) \cdot X$ for all stopping times T; more generally, $K \cdot (H \cdot X) = (HK) \cdot X$.

For $\mathscr{H}^2_{\mathrm{loc}}$, we denote by $L^2(X)$ (resp. $L^2_{\mathrm{loc}}(X)$) the set of all predictable processes H such that the process $H^2 \cdot \langle X, X\rangle$ is integrable (resp. locally integrable). Since $\langle X, X\rangle \in \mathscr{A}^+_{\mathrm{loc}}$, all locally bounded predictable processes belong to $L^2_{\mathrm{loc}}(X)$.

Theorem 1.10 *Let $X \in \mathscr{H}^2_{loc}$. The map $H \to H \cdot X$ (defined either on $\mathscr{E}$ by (3) or for all locally bounded predictable H by Theorem 1.9 has a further extension to the set $L^2_{loc}(X)$, still denoted by $H \to H \cdot X$, which meets i), ii) in Theorem 1.9, and if a sequence (H^n) of predictable processes converges pointwise to a limit H, and $|H^n| \leq K$ for some $K \in L^2_{loc}(X)$, then $sup_{s\leq t}|H^n \cdot X_s - H \cdot X_s| \xrightarrow{P} 0$ for all $t \in \mathbb{R}_+$. Moreover this extension is unique (up to evanescence), and we have:*

1. *$H \cdot X \in \mathscr{H}^2_{loc}$.*
2. *$H \cdot X \in \mathscr{H}^2$ if and only if $H \in L^2(X)$.*
3. *Properties (P1), (P3), (P4), (P5) (for $H \in L^2_{loc}(X)$ and $K \in L^2_{loc}(H \cdot X)$) hold.*
4. *If $X, Y \in \mathscr{H}^2_{loc}$ and $H \in L^2_{loc}(X)$ and $K \in L^2_{loc}(Y)$, then*
$$\langle H \cdot X, K \cdot Y\rangle = (HK) \cdot \langle X, Y\rangle.$$

We will now show that the stochastic integral of a predictable process that is càg may be approximated by Riemann sums. We call an *adapted subdivision* any sequence $\tau = (T_n)_{n\in\mathbb{N}}$ of stopping times with $T_0 = 0$, $\sup_n T_n < \infty$, and $T_n < T_{n+1}$ on $\{T < \infty\}$ (a *deterministic* subdivision if all the T_n's are constant). The *τ-Riemann approximant* of $H \cdot X$ is the process $\tau(H \cdot X)$ defined by

$$\tau(H \cdot X)_t = \sum_{n\in\mathbb{N}} H_{T_n}(X_{T_{n+1}\wedge t} - X_{T_n\wedge t}).$$

A sequence $(\tau_n = (T(n,m))_{m\in\mathbb{N}})_{n\in\mathbb{N}}$ of adapted subdivisions is called a *Riemann sequence* if $\sup_{m\in\mathbb{N}}[T(n,m+1)\wedge t - T(n,m)\wedge t] \to 0$ for all $t \in \mathbb{R}_+$ (that is, if the mesh of the restriction of the subdivisions τ_n to each interval $[0,t]$ tends to 0).

Proposition 1.13 *Let X be a semimartingale, H be a càg adapted process, and (τ_n) a Riemann sequence of adapted subdivisions. Then the τ_n-Riemann approximants $\tau_n(H \cdot X)$ converge to $H \cdot X$ in measure, uniformly on each compact interval.*

1.5 Quadratic Variation, Ito's Formula

In this section we state Ito's formula for semimartingales. The following result gives an explicit solution to a "stochastic differential equation" and describes some of its properties.

Let X and Y be two semimartingales. The *quadratic co- variation* of X and Y, denoted by $[X, Y]$ (the *quadratic variation* of X, when $Y = X$) is the process defined by:

$$[X, Y] = XY - X_0Y_0 - X_- \cdot Y - Y_- \cdot X$$

(it is defined uniquely, up to an evanescent set). Note that the following properties hold:

$$[X, Y]_0 = 0, \quad [X, Y] = [X - X_0, Y - Y_0],$$

$$[X, Y] = \frac{1}{4}([X + Y, X + Y] - [X - Y, X - Y]). \tag{4}$$

Theorem 1.11 *Let X and Y be two semimartingales.*

a) For any Riemann sequence $\{\tau_n = (T(n, m))_{m\in\mathbb{N}}\}_{n\in\mathbb{N}}$ of adapted subdivisions, the processes $S_{\tau_n}(X, Y)$ defined by

$$S_{\tau_n}(X, Y)_t = \sum_{m\geq 1}(X_{T(n,m+1)\wedge t} - X_{T(n,m)\wedge t})(Y_{T(n,m+1)\wedge t} - Y_{T(n,m)\wedge t})$$

converge to the process $[X, Y]$ in measure, uniformly on every compact interval.

b) $[X, Y] \in \mathscr{V}$ and $[X, X] \in \mathscr{V}^+$.

c) $\Delta[X, Y] = \Delta X \Delta Y$.

Let $X \in \mathscr{H}^2$. It follows from the definition of $[X, Y]$ that $X^2 - X_0^2 - [X, X] \in \mathscr{L}$. By Theorem 1.7, we have that $X^2 - X_0^2 - \langle X, X\rangle \in \mathscr{L}$. It follows that $[X, X] - \langle X, X\rangle \in \mathscr{L} \cap \mathscr{V}$ so that $[X, X] \in \mathscr{A}_{\text{loc}}$ and $\langle X, X\rangle$ is the compensator of $[X, X]$. The argument extends to $[X, Y]$ with $Y \neq X$ by (4). We formalize this result as follows:

Proposition 1.14 *Let X and Y be two local martingales.*

a) $XY - X_0Y_0 - [X, Y]$ is a local martingale.

b) If $X, Y \in \mathscr{H}^2_{loc}$ then $[X, Y] \in \mathscr{A}_{loc}$ and its compensator is $\langle X, Y\rangle$; if moreover $X, Y \in \mathscr{H}^2$, then $XY - [X, Y] \in \mathscr{M}$.

c) X belongs to $\mathscr{H}^2$ (resp. $\mathscr{H}^2_{loc}$) if and only if $[X, X]$ belongs to $\mathscr{A}$ (resp. $\mathscr{A}_{loc}$) and X_0 is square-integrable.

d) $X = X_0$ *a.s. if and only if* $[X, X] = 0$.

Theorem 1.12 *Let* X *and* Y *be two semimartingales and denote by* X^c, Y^c *their continuous martingale parts. Then*

$$[X, Y]_t = \langle X^c, Y^c \rangle_t + \sum_{s \leq t} \Delta X_s \Delta Y_s. \tag{5}$$

Theorem 1.13 (Ito's formula) *Let* $X = (X^1, \ldots, X^d)$ *be a d-dimensional semimartingale, and let* f *be a class* C^2 *function on* $\mathbb{R}^d$. *Then* $f(X)$ *is a semimartingale and we have:*

$$\begin{aligned} f(X_t) = f(X_0) &+ \sum_{i \leq d} D_i f(X_-) \cdot X^i + \frac{1}{2} \sum_{i,j \leq d} D_{ij} f(X_-) \cdot \langle X^{i,c}, X^{j,c} \rangle \\ &+ \sum_{s \leq t} \left[f(X_s) - f(X_{s-}) - \sum_{i \leq d} D_i f(X_{s-}) \Delta X_s^i \right], \end{aligned} \tag{6}$$

where $D_i f$ *and* $D_{ij} f$ *denote the partial derivatives* $\partial f / \partial x^i$ *and* $\partial^2 f / \partial x^i \partial x^j$.

As an application of Ito's formula, we will study the equation

$$Y = 1 + Y_- \cdot X \quad (\text{or:}\, dY = Y_- dX \text{ and } Y_0 = 1)$$

where X is a given semimartingale and Y is an unknown càdlàg adapted process. We will consider X to be a complex-valued semimartingale, that is, $X = X' + iX''$ with X' and X'' two real-valued semimartingales. Then (6) is read as a system of equations with real-valued terms as follows:

$$\begin{aligned} Y' &= 1 + Y'_- \cdot X' - Y''_- \cdot X'' \\ Y'' &= \qquad Y''_- \cdot X' - Y'_- \cdot X'' \end{aligned}$$

and $Y = Y' + iY''$.

Theorem 1.14 *Let* $X = X' + iX''$ *be a complex-valued semimartingale. Then equation (6) has one and only one (up to indistinguishability) càdlàg adapted solution. This solution is a semimartingale, is denoted by* $\mathscr{E}(X)$, *and is given by*

$$\begin{aligned} \mathscr{E}(X)_t = \{\exp(X_t - X_0 - \frac{1}{2}\langle X'^c, X'^c \rangle_t + \frac{1}{2}\langle X''^c, X''^c \rangle_t - i \langle X'^c, X''^c \rangle_t)\} \\ \times \prod_{s \leq t} [(1 + \Delta X_s) e^{-\Delta X_s}] \end{aligned}$$

where the (possibly infinite) product is absolutely convergent. Furthermore,

a) If X has finite variation, then so has $\mathscr{E}(X)$.

b) If X is a local martingale, then so is $\mathscr{E}(X)$.

c) Let $T = \inf(t\colon \Delta X_t = -1)$. Then $\mathscr{E}(X) \neq 0$ on the interval $[\![0, T[\![$, and $\mathscr{E}(X)_- \neq 0$ on the interval $[\![0, T]\!]$, and $\mathscr{E}(X) = 0$ on the interval $[\![T, \infty[\![$.

In particular, (5) implies that when X has finite variation, then

$$\mathscr{E}(X)_t = e^{X_t - X_0} \prod_{s \leq t} (1 + \Delta X_s) e^{-\Delta X_s}.$$

When X is a real-valued semimartingale, then

$$\mathscr{E}(X)_t = e^{X_t - X_0 - \frac{1}{2}\langle X^c, X^c \rangle_t} \prod_{s \leq t} (1 + \Delta X_s) e^{-\Delta X_s}.$$

2 Characteristics of Semimartingales

Our goal in this section is to describe a convenient way of representing semimartingales. To this end, we define a set of "characteristics" associated with a semimartingale X. We then define a "canonical representation", which expresses X in terms of its characteristics. Section 2.1 is an outline of the results concerning random measures that will be needed.

Let $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_t)_{t \geq 0}, P)$ be a stochastic basis. In addition, let $(E, \mathscr{E})$ be an auxiliary measurable space which we stipulate to be a *Blackwell space*. In this paper, however, we will assume the special case $E = \mathbb{R}^d$ with its Borel σ-field.

2.1 Random Measures

In order to characterize the behavior of the jumps of the processes of interest, we introduce the concept of an integer-valued random measure and its compensator. We also define an integral with respect to these. Then, after introducing some necessary notation, we define an integral with respect to the "compensated" random measure. This important result is closely related to the purely discontinuous part of a local martingale which was introduced in the previous section.

A *random measure* on $\mathbb{R}_+ \times E$ is a family $\mu = \{\mu(\omega; dt, dx) : \omega \in \Omega\}$ of nonnegative measures on $(\mathbb{R}_+ \times E, \mathscr{R}_+ \otimes \mathscr{E})$ such that $\mu(\omega; \{0\} \times E) = 0$ identically.

We will be using the following notation:

$\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$, with σ-fields $\tilde{\mathscr{O}} = \mathscr{O} \otimes \mathscr{E}$ and $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{E}$. Here $\mathscr{O}$ and $\mathscr{P}$ are the optional and predictable σ-fields on $(\Omega \times \mathbb{R}_+)$, respectively.

A function W on $\tilde{\Omega}$ that is $\tilde{\mathscr{O}}$- measurable(resp. $\tilde{\mathscr{P}}$-measurable) is called an *optional*(resp. a *predictable*) function.

Let μ be a random measure and W an optional function on $\tilde{\Omega}$. Define the *integral process* $W * \mu$ by

$$W * \mu(\omega) = \int_{[0,t]\times E} W(\omega, s, x)\mu(\omega; ds, dx)$$

if $\int_{[0,t]\times E} |W(\omega, s, x)|\mu(\omega; ds, dx) < \infty$.

A random measure μ is called *optional*(resp. *predictable*) if the process $W * \mu$ is optional(resp. predictable) for every optional(resp. predictable) function W.

An optional measure μ is called *integrable* if the random variable
$1 * \mu_\infty = \mu(\cdot, \mathbb{R}_+ \times E)$ is integrable (or equivalently, if $1 * \mu \in \mathscr{A}^+$).

An optional random measure μ is called $\tilde{\mathscr{P}}$-σ- *finite* if there exists a strictly positive predictable function V on $\tilde{\Omega}$ such that the random variable $V * \mu_\infty$ is integrable(or equivalently, if $V * \mu \in \mathscr{A}^+$). This property is equivalent to the existence of a $\tilde{\mathscr{P}}$-measurable partition (A_n) of $\tilde{\Omega}$ such that each $(1_{A_n} * \mu)_\infty$ is integrable.

Theorem 2.1 *Let μ be an optional $\tilde{\mathscr{P}}$-σ-finite random measure. There exists a random measure, called the* compensator *of μ and denoted by μ^p, which is unique up to a P-null set, and which is characterized as being a predictable random measure satisfying either one of the two following equivalent properties:*

1. *$E(W * \mu^p_\infty) = E(W * \mu_\infty)$ for every nonnegative $\tilde{\mathscr{P}}$-measurable function W on $\tilde{\Omega}$.*
2. *For every $\tilde{\mathscr{P}}$-measurable function W on $\tilde{\Omega}$ such that $|W| * \mu \in \mathscr{A}^+_{loc}$, $|W| * \mu^p \in \mathscr{A}^+_{loc}$, and $W * \mu - W * \mu^p$ is a local martingale.*

An *integer-valued random measure* is a random measure that satisfies:

1. $\mu(\omega; \{t\} \times E) \leq 1$ identically,
2. for each $A \in \mathscr{R}_+ \otimes \mathscr{E}$, $\mu(\cdot, A)$ takes its values in $\overline{\mathbb{N}}$;
3. μ is optional and $\tilde{\mathscr{P}}$-σ-finite.

Proposition 2.1 *If μ is an integer-valued random measure, there exists a thin random set D and an E-valued optional process β such that*

$$\mu(\omega; dt, dx) = \sum_{s \geq 0} 1_D(\omega, s)\varepsilon_{(s,\beta_s(\omega))}(dt, dx), \tag{7}$$

where ε_a denotes the Dirac measure at the point a.

Note that if (T_n) is a sequence of stopping times that exhausts the thin set D, then

$$\begin{aligned} W * \mu_t &= \sum_n W(T_n, \beta_{T_n}) 1_{\{T_n \leq t\}} \\ &= \sum_{0<s\leq t} W(s, \beta_s) 1_D(s). \end{aligned}$$

where W is any nonnegative optional function.

Proposition 2.2 *Let X be an adapted càdlàg $\mathbb{R}^d$-valued process. Then*

$$\mu^X(\omega; dt, dx) = \sum_s 1_{\{\Delta X_s(\omega)\neq 0\}}\varepsilon_{(s,\Delta X_s(\omega))}(dt, dx) \tag{8}$$

defines an integer-valued random measure on $\mathbb{R}_+ \times \mathbb{R}^d$. (In the representation (7) we have $D = \{\Delta X \neq 0\}$ and $\beta = \Delta X$). We call μ^X the random measure associated with the jumps of *X.*

Proposition 2.3 *Let μ be an integer-valued random measure, $\nu = \mu^p$ its compensator, and $J = \{(\omega, t) : \nu(\omega; \{t\} \times E) > 0\}$.*

i) J is the predictable support *$\{^p(1_D)\}$ of the set D in (7), and for all predictable times T and nonnegative predictable W:*

$$\int_E W(T, x)\nu(\{T\} \times dx) = E[W(T, \beta_T)1_D(T)|\mathscr{F}_{T-}] \ \ \text{on } \{T < \infty\}.$$

ii) There is a version of ν such that $\nu(\omega; \{t\} \times E) \leq 1$ identically and that the thin set J is exhausted by a sequence of predictable times.

Example (Poisson Measures). An *extended Poisson measure* on $\mathbb{R}_+ \times E$, relative to the filtration $\mathbf{F}$, is an integer-valued random measure μ such that

(i) the positive measure m on $\mathbb{R}_+ \times E$ defined by $m(A) = \mathrm{E}[\mu(A)]$ is σ-finite;

(ii) for every $s \in \mathbb{R}_+$ and every $A \in \mathscr{R}_+ \otimes \mathscr{E}$ such that $A \subset (s, \infty) \times \mathscr{E}$ and that $m(A) < \infty$, the variable $\mu(\cdot, A)$ is independent of the σ-field $\mathscr{F}_s$.

The measure m is called the *intensity measure* of μ. If m satisfies $m(\{t\}\times E) = 0$ for each $t \in \mathbb{R}_+$, then μ is called a *Poisson measure*; if m has the form $m(dt, dx) = dt \times F(dx)$, where F is a positive σ-finite measure on $(E, \mathscr{E})$, then μ is called a *homogeneous Poisson measure.*

Proposition 2.4 *Let μ be an extended Poisson measure on $\mathbb{R}_+ \times E$, relative to the filtration $\mathbf{F}$, with intensity measure m. Then its compensator is $\mu^p(\omega; \cdot) = m(\cdot)$.*

Let μ be an integer-valued random measure on $\mathbb{R}_+ \times E$. Let ν be a version of the dual predictable projection given in Proposition 2.3. We will now define stochastic integrals with respect to the "compensated" integer-valued random measure $(\mu - \nu)$. We will use the following notation:

$$\begin{aligned} a_t(\omega) &= \nu(\omega; \{t\} \times E), \\ J &= \{a > 0\}, \text{ exhausted by the sequence } (T_n) \\ &\qquad\qquad\text{of predictable times,} \\ \nu^c(\omega; dt, dx) &= 1_{J^c}(\omega, t). \end{aligned}$$

For any measurable function W on $\tilde{\Omega}$ we define the process $\hat{W}$ by

$$\hat{W}_t(\omega) = \int_E W(\omega, t, x)\nu(\omega; \{t\} \times dx)$$

if $\int_E |W(\omega, t, x)|\nu(\omega; \{t\} \times dx) < \infty$.

Lemma 2.1 *If W is $\tilde{\mathscr{P}}$-measurable, then $\hat{W}$ is predictable and it is a version of the predictable projection of the process $(\omega, t) \to W(\omega, t, \beta_t(\omega))1_D(\omega, t)$. In particular, for all predictable times T,*

$$\hat{W}_t = E[W(T, \beta_T)1_D(T)|\mathscr{F}_{T-}] \;\; on \; \{T < \infty\}.$$

We denote by $G_{\text{loc}}(\mu)$ the set of all $\tilde{\mathscr{P}}$-measurable real-valued functions W on $\tilde{\Omega}$ such that the process $\tilde{W}$ defined by

$$\tilde{W}_t(\omega) = W(\omega, t, \beta_t(\omega))1_D(\omega, t) - \hat{W}_t(\omega)$$

satisfies $\left[\sum_{s\leq \cdot}(\tilde{W}_s)^2\right]^{1/2} \in \mathscr{A}^+_{\text{loc}}$.

If $W \in G_{\text{loc}}(\mu)$ we call *stochastic integral of W with respect to $\mu - \nu$*, denoted by $W * (\mu - \nu)$, any purely discontinuous local martingale X such that ΔX and $\tilde{W}$ are indistinguishable.

Proposition 2.5 *Let W be a predictable function on $\tilde{\Omega}$ such that $|W|*\mu \in \mathscr{A}^+_{loc}$ (or equivalently, $|W|*\nu \in \mathscr{A}^+_{loc}$). Then $W \in G_{loc}(\mu)$ and*

$$W*(\mu-\nu) = W*\mu - W*\nu.$$

We now characterize the property $W \in G_{\text{loc}}(\mu)$ by the integrability of a suitable increasing predictable process. To this end, we associate to any predictable function W on $\tilde{\Omega}$ two increasing (possibly infinite) predictable processes as follows:

$$\begin{aligned} C(W)_t &= (W-\hat{W})^2 * \nu_t + \sum_{s\le t}(1-a_s)(\hat{W}_s)^2, \\ \overline{C}(W)_t &= |W-\hat{W}| * \nu_t + \sum_{s\le t}(1-a_s)|\hat{W}_s|. \end{aligned}$$

Theorem 2.2 *Let W be a predictable function on $\tilde{\Omega}$.*

a) W belongs to $G_{loc}(\mu)$ and $W*(\mu-\nu)$ belongs to $\mathscr{H}^2$ (resp. $\mathscr{H}^2_{loc}$) if and only if $C(W)$ belongs to $\mathscr{A}^+$ (resp. $\mathscr{A}^+_{loc}$), in which case

$$\langle W*(\mu-\nu), W*(\mu-\nu)\rangle = C(W).$$

b) W belongs to $G_{loc}(\mu)$ and $W*(\mu-\nu)$ belongs to $\mathscr{A}$ (resp. $\mathscr{A}_{loc}$) if and only if $\overline{C}(W)$ belongs to $\mathscr{A}^+$ (resp. $\mathscr{A}^+_{loc}$).

c) W belongs to $G_{loc}(\mu)$ if and only if $C(W')+\overline{C}(W'')$ belongs to $\mathscr{A}_{loc}$, where

$$\begin{aligned} W' &= (W-\hat{W})1_{\{|W-\hat{W}|\le 1\}} + \hat{W}1_{\{|\hat{W}|\le 1\}}, \\ W'' &= (W-\hat{W})1_{\{|W-\hat{W}|> 1\}} + \hat{W}1_{\{|\hat{W}|> 1\}}. \end{aligned}$$

d) Assume in addition that $\tilde{W} \ge -1$ identically. Then $\hat{W} \le 1$ on $\{a<1\}$ up to an evanescent set, and W belongs to $G_{loc}(\mu)$ if and only if the increasing predictable process $C'(W)$ defined by

$$C'(W) = \left(1-\sqrt{1+W-\hat{W}}\right)^2 * \nu_t + \sum_{s\le t}(1-a_s)\left(1-\sqrt{1-\hat{W}}\right)^2$$

belongs to $\mathscr{A}^+_{loc}$.

2.2 Definition of the Characteristics, Canonical Representation

The characteristics of a semimartingale which we present in this section can be interpreted as an extension of the terms that characterize the distribution

of a process with independent increments. Frequently, statements about semimartingales will be stated in terms of their characteristics.

Consider a d-dimensional semimartingale $X = (X^1, \ldots, X^d)$ defined on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$. Denote by $\mathscr{C}_{\mathrm{t}}^d$ the class of functions $h : \mathbb{R}^d \to \mathbb{R}^d$ which are bounded, have compact support, and satisfy $h(x) = x$ in a neighborhood of 0. In everything that follows, it may be assumed that $h(x) = x1_{\{|x|\leq 1\}}$.

Fix a truncation function $h \in \mathscr{C}_{\mathrm{t}}^d$. The process $X(h)$ defined by

$$X(h) = X - \sum_{s\leq t}[\Delta X_s - h(\Delta X_s)]$$

is a special semimartingale, and we consider its canonical decomposition

$$X(h) = X_0 + M(h) + B(h), \quad M(h) \in \mathscr{L}^d, \quad B(h) \in \mathscr{V}^d \text{ predictable.} \tag{9}$$

The *characteristics of X associated with h* is the triplet (B, C, ν) consisting in:

1. $B = (B^i)_{i\leq d}$ is the predictable process $B = B(h) \in \mathscr{V}^d$ appearing in (9) above.
2. $C = (C^{ij})_{i,j\leq d}$ is the continuous process $C^{ij} = \langle X^{i,c}, X^{j,c}\rangle \in \mathscr{V}^{d\times d}$.
3. ν is the compensator of the random measure μ^X associated with the jumps of X.

Proposition 2.6 *One can find a version of the characteristics (B, C, ν) of X which is of the form:*

$$\begin{aligned} B^i &= b^i \cdot A \\ C^{ij} &= c^{ij} \cdot A \\ \nu(\omega; dt, dx) &= dA_t(\omega)K_{\omega,t}(dx) \end{aligned}$$

where:

1. *A is a predictable process in $\mathscr{A}_{loc}^+$;*
2. *$b = (b^i)_{i\leq d}$ is a d-dimensional predictable process;*
3. *$c = (c^{ij})_{i,j\leq d}$ is a predictable process with values in the set of all symmetric nonegative $d \times d$ matrices;*
4. *$K_{\omega,t}(dx)$ is a transition kernel from $(\Omega \times \mathbb{R}_+, \mathscr{P})$ into $(\mathbb{R}^d, \mathscr{R}^d)$ which satisfies:*

$$K_{\omega,t}(\{0\}) = 0, \qquad \int K_{\omega,t}(dx)(|x|^2 \wedge 1) \leq 1,$$

$$\Delta A_t(\omega) > 0 \implies b_t(\omega) = \int K_{\omega,t}(dx)h(x),$$
$$and\ \ \Delta A_t(\omega)K_{\omega,t}(\mathbb{R}^d) \le 1.$$

The following result characterizes semimartingales with independent increments in terms of the characteristics.

Theorem 2.3 *Let X be a d-dimensional semimartingale with $X_0 = 0$. Then it is a process with independent increments if and only if there is a version (B, C, ν) of its characteristics that is deterministic.*

Proposition 2.7 *Let X be a semimartingale with characteristics $(B(h), C, \nu)$ relative to a truncation function h. X is a special semimartingale if and only if $(|x|^2 \wedge |x|) * \nu \in \mathscr{A}_{loc}$. In this case, the canonical decomposition $X = X_0 + N + A$ satisfies $A = B(h) + (x - h(x)) * \nu$.*

Theorem 2.4 *Let X be a d-dimensional semimartingale, with characteristics (B, C, ν) relative to a truncation function $h \in \mathscr{C}_t^d$, and with the measure μ^X associated to its jumps by (8). Then $W^i(\omega, t, x) = h^i(x)$ belongs to $G_{loc}(\mu^X)$ for all $i \le d$, and the following canonical representation of X holds:*

$$X = X_0 + X^c + h * (\mu^X - \nu) + (x - h(x)) * \mu^X + B.$$

The following corollary follows from the last two results and Proposition 2.5. It implies that in the case of a special semimartingale we can take $h(x) = x$.

Corollary 2.1 *Let X be a d-dimensional special semimartingale with characteristics (B, C, ν) and μ^X the measure associated to its jumps. Then $W^i(\omega, t, x) = x^i$ belongs to $G_{loc}(\mu^X)$, and if $X = X_0 + N + A$ is its canonical decomposition, then*

$$X = X_0 + X^c + x * (\mu^X - \nu) + A.$$

3 Martingale Problems, Diffusion Processes

In this section we present diffusion processes, which will be used to construct a term structure model in Section 6. First we introduce the notion of a martingale problem, which describes a useful framework in which to characterize the set of probability measures under which a suitable process is a semimartingale. The connection between the two concepts is established in Theorem 3.2.

We will be working in the following setting. Let $(\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_t)_{t \ge 0})$ be a filtered space, and let $\mathscr{H}$ be a sub-σ-field of $\mathscr{F}_0$, called the *initial σ-field*. Let P_H be an *initial condition*, that is, a probability measure on $(\Omega, \mathscr{H})$.

3.1 Martingale Problems

Let $X = (X^i)_{i\leq d}$ be a d-dimensional càdlàg adapted process on $(\Omega, \mathscr{F}, \mathbf{F})$. X will be a candidate for a semimartingale, so we introduce the following in relation to X.

i) $h \in \mathscr{C}_{\mathrm{t}}^d$, a truncation function;

ii) A triplet (B, C, ν) such that

  (a) $B = (B^i)_{i\leq d}$ is $\mathbf{F}$-predictable, with finite variation over finite intervals, and $B_0 = 0$;

  (b) $C = (C^{ij})_{i,j\leq d}$ is $\mathbf{F}$-predictable, continuous, $C_0 = 0$, and $C_t - C_s$ is a non-negative symmetric $d \times d$ matrix for $s \leq t$;

  (c) ν is an $\mathbf{F}$-predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$, which charges neither $\mathbb{R}_+ \times \{0\}$ nor $\{0\} \times \mathbb{R}^d$, and such that

$$(|x|^2 \wedge 1) * \nu_t(\omega) < \infty, \qquad \int \nu(\omega; \{t\} \times dx) h(x) = \Delta B_t(\omega),$$

  and $\nu(\omega; \{t\} \times \mathbb{R}^d) \leq 1$ identically.

A *solution to the martingale problem* associated with $(\mathscr{H}, X)$ and $(P_H; B, C, \nu)$ is a probability measure P on $(\Omega, \mathscr{F})$ such that:

1. the restriction $P_{|\mathscr{H}}$ of P to $\mathscr{H}$ equals P_H;
2. X is a semimartingale on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with characteristics (B, C, ν) relative to the truncation function h.

The set of all solutions P will be denoted by $\mathfrak{s}(\mathscr{H}, X | P_H; B, C, \nu)$.

Sometimes we will impose additional structure on $(\Omega, \mathscr{F}, \mathbf{F})$ as follows:

1. $\mathbf{F}$ is *generated* by X and $\mathscr{H}$, by which we mean:
   (a) $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s^0$ and $\mathscr{F}_s^0 = \mathscr{H} \vee \sigma(X_r : r \leq s)$ (i.e. $\mathbf{F}$ is the smallest filtration such that X is adapted and $\mathscr{H} \subset \mathscr{F}_0$);
   (b) $\mathscr{F} = \mathscr{F}_{\infty-}$ $(= \vee_t \mathscr{F}_0)$.
2. (*The canonical setting*) Ω is the *canonical space* of càdlàg functions $\omega : \mathbb{R}_+ \to \mathbb{R}^d$; X is the *canonical process* defined by $X_t(\omega) = \omega(t)$; $\mathscr{H} = \sigma(X_0)$; $\mathbf{F}$ is generated by X and $\mathscr{H}$ in the sense of 1. above.

When $\mathscr{H} = \sigma(X_0)$, we can identify the initial measure P_H with the distribution of X_0 as follows: If η is a probability measure on $\mathbb{R}^d$, we also denote by η the measure on $(\Omega, \mathscr{H})$ defined by $\eta(X_0 \in A) = \eta(A)$.

The next result partly justifies the restrictions introduced above.

Theorem 3.1 *Let (B, C, ν) meet ii) above and be deterministic.*

a) If $\mathscr{F}$ is generated by X and $\mathscr{H}$ then $\mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$ contains at most one element P.

b) Under the canonical setting, for any probability measure η on $\mathbb{R}^d$, $\mathfrak{s}(\mathscr{H}, X|\eta; B, C, \nu)$ contains one and only one solution.

3.2 Diffusion Processes

We now assume the canonical setting defined above. Let P be a probability measure on $(\Omega, \mathscr{F})$. X is called a *diffusion process with jumps* on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ if it is a semimartingale with the following characteristics (the truncation function h is fixed):

$$\begin{aligned} B_t^i(\omega) &= \int_0^t b^i(s, X_s(\omega))ds \quad (= +\infty \text{ if integral diverges}) \\ C_t^{ij}(\omega) &= \int_0^t c^{ij}(s, X_s(\omega))ds \quad (= +\infty \text{ if integral diverges}) \qquad (10) \\ \nu(\omega; dt \times dx) &= dt \times K_t(X_t(\omega), dx), \end{aligned}$$

where

$$\begin{aligned} &b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \text{ is Borel} \\ &c : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d \text{ is Borel, } c(s,x) \text{ is symmetric nonnegative} \\ &K_t(x, dy) \text{ is a Borel transition kernel from } \mathbb{R}_+ \times \mathbb{R}^d \text{ into } \mathbb{R}^d, \\ &\qquad \text{with } K_t(x, \{0\}) = 0. \end{aligned}$$

Moreover,
a) if $\nu = 0$, X is called a *diffusion* (it is then a.s. continuous);
b) if $b(s,x)$, $c(s,x)$, $K_s(x,dy)$ do not depend upon s, X is called a *homogeneous diffusion* (with jumps).

We now introduce some notation in order to define a stochastic differential equation related to a diffusion. Let $\mathscr{B}' = (\Omega', \mathscr{F}', \mathbf{F}', P')$ be another stochastic basis endowed with the following *driving terms*:

1. $W = (W^i)_{i \le m}$, an m-dimensional standard Wiener process

2. $\mathfrak{p}$, a Poisson random measure on $\mathbb{R}_+ \times E$ with intensity measure $q(dt, dx) = dt \otimes F(dx)$. Here $(E, \mathscr{E})$ is a measurable space as in Section 2 (recall that we assume $E = \mathbb{R}^d$), and F is a positive σ-finite measure on $(E, \mathscr{E})$.

We assume that the following coefficients are given:

$$\begin{aligned} &\beta = (\beta^i)_{i\le d}, \text{ a Borel function: } \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \\ &\gamma = (\gamma^{ij})_{i\le d, j\le m}, \text{ a Borel function: } \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^m \\ &\delta = (\delta^i)_{i\le d}, \text{ a Borel function: } \mathbb{R}_+ \times \mathbb{R}^d \times E \to \mathbb{R}^d. \end{aligned} \tag{11}$$

We also let ξ be a given $\mathscr{F}'_0$-measurable $\mathbb{R}^d$- valued random variable which we call the *initial variable*.

Define a stochastic differential equation as follows:

$$\begin{aligned} dY_t =& \beta(t, Y_t)dt + \gamma(t, Y_t)dW_t + h \circ \delta(t, Y_{t-}, z)(\mathfrak{p}(dt, dz) - q(dt, dz)) \\ &+ h' \circ \delta(t, Y_{t-}, z)\mathfrak{p}(dt, dz) \end{aligned} \tag{12}$$

with $Y_0 = \xi$. Here h is a truncaton function and $h'(x) = x - h(x)$.

A *solution-process*(*strong solution*) to (12), on the basis $\mathscr{B}'$ and relative to the driving terms $(W, \mathfrak{p})$, is a càdlàg adapted process Y such that for each $i \le d$,

$$\begin{aligned} Y^i = \xi^i + &\beta^i(Y) \cdot t + \sum_{j\le m} \gamma^{ij}(Y_-) \cdot W^j + h^i \circ \delta(Y_-) * (\mathfrak{p} - q) \\ &+ h'^i \circ \delta(Y_-) * \mathfrak{p}. \end{aligned} \tag{13}$$

A *solution-measure* (or *weak solution*) to (12) with initial condition η (a probability measure on $\mathbb{R}^d$) is a probability measure P on $(\Omega, \mathscr{F})$(the canonical space) with the following property: there exists a stochastic basis $\mathscr{B}'$ with driving terms $(W, \mathfrak{p})$ and with a $\mathscr{F}'_0$-measurable variable ξ meeting $\mathcal{L}(\xi) = \eta$, and a solution-process Y on $\mathscr{B}'$, such that P be the law of Y.

Note that if Y is a solution-process then the above expression gives the canonical representation of Y:

$$Y_0 = \xi, \qquad Y^{i,c} = \sum_{j\le m} \gamma^{ij}(Y_-) \cdot W^j, \qquad h * (\mu^Y - \nu) = h \circ \delta(Y_-) * (\mathfrak{p} - q)$$

$$(x - h(x)) * \mu^Y = h' \circ \delta(Y_-) * \mathfrak{p}, \qquad B = \beta(Y) \cdot t.$$

Theorem 3.2 *Let η be an initial condition (a probability on $\mathbb{R}^d$), and β, γ, δ be coefficients as in (11). The set of all solution-measures to (12) with initial condition η is the set $\mathfrak{s}(\mathscr{H}, X|\eta; B, C, \nu)$ of all solutions to a martingale problem on the canonical space, where (B, C, ν) are given by (10) with*

$$b = \beta, \qquad c = \gamma\gamma^T \qquad \left(i.e.\ c^{ij} = \sum_{1\leq k\leq m} \gamma^{ik}\gamma^{jk}\right)$$

$$K_t(y, A) = \int 1_{A\setminus\{0\}}\left(\delta(t, y, z)\right) F(dz).$$

Theorem 3.3 *Assume the following two conditions:*

1. Local Lipshitz coefficients. *For each $n \in \mathbb{N}^*$ there is a constant θ_n and a function $\rho_n\colon E \to \mathbb{R}_+$ with $\int \rho_n(z)^2 F(dz) < \infty$, such that for $t \leq n$, $|y| \leq n$, $|y'| \leq n$:*

$$|\beta(s, y) - \beta(s, y')| \leq \theta_n|y - y'|, \quad |\gamma(s, y) - \gamma(s, y')| \leq \theta_n|y - y'|,$$

$$|h \circ \delta(s, y, z) - h \circ \delta(s, y', z)| \leq \rho_n(z)|y - y'|,$$
$$|h' \circ \delta(s, y, z) - h' \circ \delta(s, y', z)| \leq \rho_n(z)^2|y - y'|.$$

2. Linear growth. *For each $n \in \mathbb{N}^*$ there are θ_n and ρ_n as above, such that for all $t \leq n$ and all $y \in \mathbb{R}^d$:*

$$|\beta(s, y)| \leq \theta_n(1 + |y|), \quad |\gamma(s, y)| \leq \theta_n(1 + |y|),$$

$$|h \circ \delta(s, y, z)| \leq \rho_n(z)|y - y'|,$$
$$|h' \circ \delta(s, y, z)| \leq [\rho_n(z)^2 \wedge \rho_n(z)^4](1 + |y|).$$

Then (12) has a solution-process Y, and only one (up to indistinguishability) on any stochastic basis $\mathscr{B}'$ supporting driving terms $(W, \mathfrak{p})$.

Theorem 3.4 *Suppose that on any stochastic basis $\mathscr{B}'$ supporting driving terms $(W, \mathfrak{p})$, there is at most one solution-process (up to indistinguishability). Then, if there is a solution-measure, with a given initial condition, this solution- measure is unique.*

4 Changes of Measures

Our goal in this section is to characterize the behavior of a semimartingale after a change of probability measure. In particular, assume X is a P-semimartingale and P' is another probability measure such that $P' \overset{\text{loc}}{\ll} P$. Then Girsanov's theorem states that X is also a P'-semimartingale and gives its P'-characteristics

in terms of its P-characteristics. These results have important applications in finance as will be seen in Section 6.

The setting is the same as in Section 2. measurable space $(E, \mathscr{E})$. To every random measure μ on $\mathbb{R}_+ \times E$ defined on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ we associate the following:

M_μ^P is the positive measure on $(\tilde{\Omega}, \mathscr{F} \otimes \mathscr{R}_+ \otimes \mathscr{E})$ defined by $M_\mu^P(W) = \mathrm{E}(W * \mu_\infty)$ for all measurable nonnegative functions W.

Assume that μ is $\tilde{\mathscr{P}}$-σ-finite on $(\Omega, \mathscr{F}, \mathbf{F}, P)$. For every nonnegative measurable function W we call the *conditional expectation relative to* M_μ^P with respect to $\tilde{\mathscr{P}}$ the M_μ^P- a.e unique $\tilde{\mathscr{P}}$-measurable function $W' = M_\mu^P(W|\tilde{\mathscr{P}})$ such that

$$M_\mu^P(WU) = M_\mu^P(W'U)$$

for all nonnegative $\tilde{\mathscr{P}}$-measurable U.

Theorem 4.1 (Girsanov's Theorem for Random Measures) *Assume that $P' \overset{loc}{\ll} P$ and let Z be the density process. Let μ be an integer-valued random measure on $\mathbb{R}_+ \times E$ defined on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ (this implies in particular that it is $\tilde{\mathscr{P}}$- σ-finite relative to P), and denote by ν its P-compensator.*

a) *μ is also $\tilde{\mathscr{P}}$-σ-finite relative to P'.*

b) *Let Y be a $\tilde{\mathscr{P}}$-measurable nonnegative function on $\tilde{\Omega}$. Let ν' be a version of the P'-compensator of μ. There is equivalence between:*

   i) *$\nu' = Y \cdot \nu$ P'-a.s. (where $Y \cdot \nu(\omega; dt, dx) = \nu(\omega; dt, dx) Y(\omega; dt, dx)$);*

   ii) *$1_{\{Z_0 > 0\}} \cdot \nu' = Y 1_{\{Z_0 > 0\}} \cdot \nu$ P-a.s.;*

   iii) *YZ_- is a version of the conditional expectation $M_\mu^P(Z|\tilde{\mathscr{P}})$.*

   Moreover, any nonnegative version Y of $M_\mu^P\left(\frac{Z}{Z_-} 1_{\{Z_0>0\}} | \tilde{\mathscr{P}}\right)$ has the above properties.

c) *There is a version of ν' that meets identically $\nu' = Y \cdot \nu$ for some $\tilde{\mathscr{P}}$-measurable nonnegative function Y.*

Theorem 4.2 (Girsanov's Theorem for Semimartingales) *Assume that $P' \overset{loc}{\ll} P$. There exist a $\tilde{\mathscr{P}}$-measurable nonnegative function Y and a predictable process $\beta = (\beta^i)_{i \le d}$ satisfying*

$$|h(x)(Y-1)| * \nu_t < \infty \quad P'\text{-a.s. for } t \in \mathbb{R}_+,$$

$$\left|\sum_{j\le d} c^{ij}\beta^j\right| \cdot A < \infty, \ and \ \left(\sum_{j,k\le d} \beta^j c^{jk}\beta^k\right) \cdot A_t < \infty \quad P'\text{-}a.s. \ for \ t \in \mathbb{R}_+,$$

and such that a version of the characteristics of X relative to P' are

$$\begin{aligned} B'^i &= B^i + \left(\sum_{j\le d} c^{ij}\beta^j\right) \cdot A + h(x)(Y-1) * \nu \\ C' &= C \\ \nu' &= Y \cdot \nu \end{aligned}$$

Moreover, Y and β meet all of the above conditions, if and only if

$$YZ_- = M^P_{\mu^X}(Z|\tilde{\mathscr{P}}) \tag{14}$$

$$\langle Z^c, X^{i,c}\rangle = \left(\sum_{j\le d} c^{ij}\beta^j Z_-\right) \cdot A \tag{15}$$

up to a P-null set, where Z is the density process, Z^c is its continuous martingale part relative to P, and $\langle Z^c, X^{i,c}\rangle$ is the bracket relative to P.

5 The Representation Property, Fundamental Representation Theorem

Let $(\Omega, \mathscr{F}, \mathbf{F}, P)$ be a stochastic basis supporting a semimartingale $X = (X^i)_{i\le d}$ with characteristics (B, C, ν) relative to a truncation function h. X^c denotes the continuous martingale part of X, and $\mu = \mu^X$ is the random measure associated with the jumps of X.

We have seen that any local martingale can be decomposed in terms of a continuous part and a purely discontinuous part. In this section we look at the structure of the decomposition of a local martingale in terms of the given semimartingale X. We first state two related results for future reference.

Proposition 5.1 *Let $M = X^c$ and Z be an arbitrary local martingale.*

a) There is a predictable process $H = (H^i)_{i\le d}$ such that

$$[Z, M^i] = \langle Z^c, M^i\rangle = \left(\sum_{j\le d} c^{ij}H^j\right) \cdot A. \tag{16}$$

b) *Any predictable process meeting (16) belongs to* $L^2_{loc}(M)$, *and the stochastic integral* $H \cdot M$ *does not depend upon the chosen version of* H, *and* $Y = Z - H \cdot M$ *is orthogonal to all components of* M *and*

$$[Y, M^i] = \langle Y^c, M^i \rangle = 0. \tag{17}$$

Proposition 5.2 *Let* N *be a local martingale, and* $U = M^P_\mu(\Delta N \mid \tilde{\mathscr{P}})$ *(here,* ΔN *is considered as defined on* $\tilde{\Omega}$ *by* $\Delta N(\omega, t, x) = \Delta N_t(\omega)$*).*

a) *There is a version of* U *such that* $\{a = 1\} \subset \{\hat{U} = 0\}$.

b) *Let* $W = U + \frac{\hat{U}}{1-a} 1_{\{a<1\}}$. *Then* $W \in G_{loc}(\mu)$, *and if* $Y = W * (\mu - \nu)$ *and* $Z = N - Y$, *then we have* $M^P_\mu(\Delta Z \mid \tilde{\mathscr{P}}) = 0$.

$M^P_\mu(\Delta Z \mid \tilde{\mathscr{P}}) = 0$ may be interpreted as Z being "orthogonal" to μ, so Y is a sort of *projection* of N on μ (or, rather, on the space of all integrals of the form $V * (\mu - \nu)$).

We say that a local martingale M has the *representation property relative to* X if it has the form

$$M = M_0 + H \cdot X^c + W * (\mu - \nu),$$

where $H = (H^i)_{i \le d}$ belongs to $L^2_{\text{loc}}(X^c)$ and $W \in G_{\text{loc}}(\mu)$.

Lemma 5.1 *Every local martingale* M *has a decomposition*

$$M = H \cdot X^c + W * (\mu - \nu) + N, \tag{18}$$

where $H \in L^2_{loc}(X^c)$, $W \in G_{loc}(\mu)$, *and*

$$\langle N^c, (X^c)^i \rangle = 0 \ \ \forall i \le d, \qquad M^P_\mu(\Delta N | \tilde{\mathscr{P}}) = 0. \tag{19}$$

Moreover, this decomposition is unique, up to indistinguishability.

Corollary 5.1 *The following statements are equivalent:*

i) *All local martingales have the representation property.*

i) *All local martingales satisfying (19) are trivial (a local martingale* N *is called* trivial *if* $N_t = N_0$ *a.s. for all* $t \in \mathbb{R}_+$*).*

iii) *All bounded martingales satisfying (19) are trivial.*

We now relate the representation property with the martingale problem. To this end, we introduce an initial condition: $\mathscr{H}$ is a sub σ-field of $\mathscr{F}_0$, and $P_H = P_{|\mathscr{H}}$ is the restriction of P to $\mathscr{H}$. Note that $P \in \mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$. The following result implies that if all local martingales have the representation property then we can say something about the uniqueness of P.

Theorem 5.1 (Fundamental Representation Theorem) *In addition to the above, assume that $\mathscr{F} = \mathscr{F}_{\infty-}$. Then the following are equivalent:*

(i) All local martingales have the representation property relative to X; moreover $\mathscr{F}_0$ is contained in the σ-field generated by $\mathscr{H}$ and the P-null sets of $\mathscr{F}$.

(ii) If $P' \in \mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$ and $P' \overset{loc}{\ll} P$, then $P' = P$.

(iii) If $P' \in \mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$ and $P' \ll P$, then $P' = P$.

Note that if $\mathscr{H} = \mathscr{F}_0$ then the second condition in (i) is met and we have the following:

Corollary 5.2 *Denote by P_0 the restriction of P to $\mathscr{F}_0$. If $\mathscr{F} = \mathscr{F}_{\infty-}$ then the following are equivalent:*

(i) All local martingales have the representation property relative to X.

(ii) If $P' \overset{loc}{\ll} P$ and $P'_0 = P_0$ and X admits (B, C, ν) for P'-characteristics, then $P' = P$.

(iii) If $P' \ll P$ and $P'_0 = P_0$ and X admits (B, C, ν) for P'- characteristics, then $P' = P$.

6 Term Structure Models

Following Björk, Kabanov and Runggaldier [2] we will outline a model for the bond price dynamics in the context of a diffusion with jumps described above.

6.1 Introduction

We assume the canonical setting. Let $\mathsf{P}(t, T)$ be the price at time t of a bond which matures at time T. It is assumed that for each $T > 0$, $(\mathsf{P}(t, T))_{0 \le t \le T}$ is an optional, $(\mathscr{F}_t)$-adapted process, and for each t, $\mathsf{P}(t, T)$ is P-a.s. continuously differentiable in the T variable. Let $f(t, T)$ denote the T-*forward rate* at time

t, defined by $f(t,T) = -\frac{\partial}{\partial T}\mathsf{P}(t,T)$. The *short rate* r is defined by $r_t = f(t,t)$, and the *money account process* B is defined by

$$\mathsf{B}_t = \exp\left(\int_0^t r_s ds\right).$$

In order to model the bond price dynamics we could start with a description of the forward rate or short rate dynamics. Alternatively, we could follow a direct approach, obtaining $\mathsf{P}(t,T)$ as the solution of a stochastic differential equation. Therefore, we are interested in studying dynamics of the following forms:

$$dr_t = a_t dt + b_t dW_t + \int_E q(t,x)\mu(dt,dx) \tag{20}$$

$$d\mathsf{P}(t,T) = \mathsf{P}(t-,T)\left\{m(t,T)dt + v(t,T)dW_t + \int_E n(t,x,T)\mu(dt,dx)\right\} \tag{21}$$

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t + \int_E \delta(t,x,T)\mu(dt,dx). \tag{22}$$

The coefficients $b(t,T)$, $v(t,T)$, and $\sigma(t,T)$ are assumed to be m- dimensional row vector processes. The following technical assumptions will be needed:

Assumption

1. For any fixed $T > 0$, $n(t,x,T)$ and $\delta(t,x,T)$ are uniformly bounded. Furthermore, for each t,

$$\int_0^t \int_E h'(n(s,x,T))F(dx)ds < \infty,$$

   where $h'(z) = |z|^2 \wedge |z|$ for $z \in \mathbb{R}$.

2. For each fixed ω, t, and (in appropriate cases) x, all the objects $m(t,T)$, $v(t,T)$, $n(t,x,T)$, $\alpha(t,T)$, $\sigma(t,T)$ and $\delta(t,x,T)$ are assumed to be continuously differentiable in the T-variable. This partial T-derivative is denoted $m_T(t,T)$, etc.

3. All processes are assumed to be regular enough to allow us to differentiate under the integral sign as well as to interchange the order of integration.

4. For any t the price curves $\mathsf{P}(\omega,t,T)$ are bounded functions for almost every ω.

Proposition 6.1 *If $f(t,T)$ satisfies (22), then $\mathsf{P}(t,T)$ satisfies*

$$d\mathsf{P}(t,T) = \mathsf{P}(t-,T)\left[\left(r_t + A(t,T) + \frac{1}{2}\|S(t,T)\|^2\right)dt + S(t,T)dW_t + \int_E \left(e^{D(t,x,T)} - 1\right)\mu(dt,dx)\right],$$

where

$$
\begin{aligned}
A(t,T) &= -\int_t^T \alpha(t,s)ds \\
S(t,T) &= -\int_t^T \sigma(t,s)ds \\
D(t,x,T) &= -\int_t^T \delta(t,x,s)ds.
\end{aligned} \tag{23}
$$

We could also introduce the models above by specifying the characteristics as in Section 3. By way of example, let F be a Lévy measure on E and define the set of characteristics (B, C, ν) by

$$
\begin{aligned}
B_t &= \int_0^t \mathsf{P}(s,T)\left(m(s,T) + \int_E n(s,x,T)F(dx)\right) ds \\
C_t &= \int_0^t \mathsf{P}(s,T)^2 v(s,T)v(s,T)^{\mathrm{T}} ds \\
\nu(dt,dx) &= \mathsf{P}(t,T)n(t,x,T)F(dx)dt.
\end{aligned}
$$

In order to express the corresponding stochastic differential equation, let $J(dt,dx)$ be the integer valued random measure constructed as in Section 2.1. We then have

$$
\frac{d\mathsf{P}(t,T)}{\mathsf{P}(t-,T)} = m(t,T)dt + v(t,T)dW_t + \int_E n(t,x,T)J(dt,dx).
$$

6.2 Example: Stable Driving Process

Let X be a Lévy process on $\mathbb{R}$ with characteristics $(tb, tc, t\nu)$, where $b \in \mathbb{R}$, $c \geq 0$ and $\int_{\mathbb{R}}(1 \wedge |x|^2)\nu(dx) < \infty$.

Let T be an increasing Lévy process on $\mathbb{R}$ with characteristics $(t\beta, 0, t\rho)$. Here $\beta \geq 0$ and $\int_{(0,\infty)}(1 \wedge \xi)\rho(d\xi) < \infty$. We call the process T a *subordinator.*

We now introduce the process $Y_t(\omega) = X_{T_t(\omega)}(\omega)$, $t > 0$, obtained by the *subordination* of X by the subordinator T. A process identical in law to Y is said to be a subordinate of X. It is a Lévy process, (see Sato [8]) with characteristics $(tb', tc', t\nu')$ where

$$
\begin{aligned}
b' &= \beta b + \int_{(0,\infty)} \rho(ds) \int_{|x|\leq 1} x P_{X_s}(dx) \\
c' &= \beta c \\
\nu'(dx) &= \beta\nu(dx) + \int_{(0,\infty)} P_{X_s}(dx)\rho(ds).
\end{aligned} \tag{24}
$$

In financial applications the subordinator T can be interpreted as the market operational time. For example, it can be used to model the arrival of news which would imply changes in market activity.

Following Hurst, Platen and Rachev(1999) we present an example of this approach using a stable subordinator. Let W be a Wiener process and let T be an $\alpha/2$-stable Lévy process such that

$$T_{t+s} - T_t \sim S_{\alpha/2}(cs^{\alpha/2}, 1, 0), \quad s, t \geq 0.$$

The parameter $\alpha/2$ denotes the index of stability. The other three parameters represent scale, skewness and location, respectively. (See Samorodnitsky and Taqqu [7]) In terms of our notation this means that the set of characteristics of T is $(0, 0, t\rho)$ where

$$\rho(d\xi) = c^{\alpha/2} 1_{(0,\infty)}(\xi) \frac{d\xi}{\xi^{1+\alpha/2}}.$$

It follows from (24) that the set of characteristics of the subordinated process W_T is $(0, 0, t\nu')$ where

$$\nu'(dx) = \frac{(2c)^{\alpha/2}}{\sqrt{\pi}} \Gamma\left(\frac{\alpha+1}{2}\right) \frac{dx}{|x|^{\alpha+1}}.$$

6.3 Bond Markets, Arbitrage

We now present the framework (Björk, Kabanov and Runggaldier [2]) in which we will state results concerning the absence of arbitrage in a model of bond prices. It will be assumed throughout that the filtration $\mathbf{F}$ is the natural filtration generated by W and μ.

A *portfolio* in the bond market is a pair (g, h), where

1. g is a predictable process.
2. For each ω, t, $h_t(\omega, \cdot)$ is a signed finite Borel measure on $[t, \infty)$.
3. For each Borel set A the process $h_t(A)$ is predictable.

The *discounted bond prices* $\overline{\mathsf{P}}(t, T)$ are defined by

$$\overline{\mathsf{P}}(t, T) = \frac{\mathsf{P}(t, T)}{\mathsf{B}_t}.$$

A portfolio (g, h) is said to be *feasible* if the following conditions hold for every t:

$$\int_0^t |g_s| ds, \qquad \int_0^t \int_s^\infty |m(s, T)||h_s(dT)| ds < \infty,$$

$$\int_0^t \int_s^\infty \int_E |n(s,x,T)||h_s(dT)|\nu(ds,dx) < \infty,$$

$$\text{and} \int_0^t \left\{ \int_s^\infty |v(s,T)||h_s(dT)| \right\}^2 ds < \infty.$$

The *value process* corresponding to a feasible portfolio $\pi = (g,h)$ is defined by

$$V_t^\pi = g_t \mathsf{B}_t + \int_t^\infty \mathsf{P}(t,T) h_t(dT).$$

The *discounted value process* is

$$\overline{V_t^\pi} = \mathsf{B}_t^{-1} V_t^\pi.$$

A feasible portfolio is said to be *admissible* if there is a number $a \geq 0$ such that $V_t^\pi \geq -a$ P-a.s. for all t.

A feasible portfolio is said to be *self-financing* if the corresponding value process satisfies

$$\begin{aligned} V_t^\pi = V_0^\pi &+ \int_0^t g_s d\mathsf{B}_s + \int_0^t \int_s^\infty m(s,t)\mathsf{P}(s,t)h_s(dT)ds \\ &+ \int_0^t \int_s^\infty v(s,t)\mathsf{P}(s,t)h_s(dT)dW_s \\ &+ \int_0^t \int_s^\infty \int_E n(s,x,T)\mathsf{P}(s-,t)h_s(dT)\mu(ds,dx). \end{aligned}$$

The preceding relation can be interpreted formally as follows:

$$dV_t^\pi = g_t d\mathsf{B}_t + \int_s^\infty h_t(dT) d\mathsf{P}(t,T).$$

A *contingent T-claim* is a random variable $X \in L_+^0(\mathscr{F}_T, P)$. An *arbitrage portfolio* is an admissible self-financing portfolio $\pi = (g,h)$ such that the corresponding value process satisfies

1. $V_0^\pi = 0$
2. $V_T^\pi \in L_+^0(\mathscr{F}_T, P)$ with $P(V_T^\pi > 0) > 0$.

If no arbitrage portfolios exist for any $T > 0$ we say that the model is *arbitrage-free*.

Take the measure P as given. We say that a positive martingale $M = (M_t)_{t\geq 0}$ with $\mathrm{E}^P[M_t] = 1$ is a *martingale density* if for every $T > 0$ the process $(\overline{\mathsf{P}}(t,T)M_t)_{0\leq t\leq T}$ ia a P- local martingale. If, moreover, $M_t > 0$ for all $t > 0$ we say that M is a *strict martingale density*.

We say that that a probability measure Q on $(\Omega, \mathscr{F})$ is a *martingale measure* if $Q_t \sim P_t$ and the process $(\overline{\mathsf{P}}(t,T))_{0\leq t\leq T}$ is a Q-local martingale for every $T > 0$. Here Q_t, P_t are the restrictions $Q_{|\mathscr{F}_t}$ and $P_{|\mathscr{F}_t}$, respectively.

Proposition 6.2 *Suppose that there exists a strict martingale density. Then the bond market model is arbitrage-free.*

We will make the following simplifying assumption:

ASSUMPTION For any positive martingale $N = (N)_t$ with $\mathrm{E}^P(N_t) = 1$ there exists a probability measure Q on $\bigcup_{t\geq 0}\mathscr{F}_t$ such that $N_t = dQ_t/dP_t$.

The following results relate the coefficients in (21) and (22) with a model free of arbitrage.

Theorem 6.1 *Let the bond price dynamics be given by (21). There exists a martingale measure if and only if the following conditions hold:*

(i) There exists a predictable process ϕ and a $\tilde{\mathscr{P}}$-measurable function $Y(\omega,t,x)$ with $Y > 0$ satisfying

$$\int_0^t \|\phi_s\|^2 ds < \infty, \qquad \int_0^t \int_E |Y(s,x) - 1| F(dx) ds < \infty.$$

and such that $E^P(\mathscr{E}(L)_t) = 1$ for all finite t, where the process L is defined by

$$L = \phi \cdot W + (Y - 1) * (\mu - \nu).$$

(ii) For all $T > 0$, and $t \in [0,T]$ we have

$$m(t,T) + \phi_t v(t,T)^T + \int_E Y(t,x) n(t,x,T) F(dx) = r_t. \tag{25}$$

The following theorem gives a similar result when we consider the forward rate dynamics.

Theorem 6.2 *Let the forward rate dynamics be given by (22). There exists a martingale measure if and only if the following conditions hold:*

(i) *There exists a predictable process* ϕ *and a* $\tilde{\mathscr{P}}$*-measurable function* $Y(\omega, t, x)$ *with* $Y > 0$ *satisfying*

$$\int_0^t \|\phi_s\|^2 ds < \infty, \qquad \int_0^t \int_E |Y(s,x) - 1| F(dx) ds < \infty.$$

and such that $E^P(\mathscr{E}(L)_t) = 1$ *for all finite* t*, where the process* L *is defined by*

$$L = \phi \cdot W + (Y - 1) * (\mu - \nu).$$

(ii) *For all* $T > 0$*, and* $t \in [0, T]$ *we have*

$$A(t,T) + \frac{1}{2}\|S(t,T)\|^2 + \phi_t S(t,T)^T + \int_E Y(t,x)\left(e^{D(t,x,T)} - 1\right) F(dx) = 0,$$

where A*,* S *and* D *are defined in (23).*

Since $n(\cdot,\cdot,T)$ in (21) is uniformly bounded, we can assume $h(x) = x$ in (13) to obtain the canonical representation

$$\frac{d\mathsf{P}(t,T)}{\mathsf{P}(t-,T)} = \left(m(t,T) + \int_E n(t,x,T)F(dx)\right) dt + v(t,T)dW_t$$
$$+ \int_E n(t,x,T)(\mu(dt,dx) - \nu(dt,dx)).$$

By Theorem 3.2 the P characteristics of $\mathsf{P}(\cdot,T)$ are

$$\begin{aligned} B_t &= \int_0^t \mathsf{P}(s,T)\left(m(s,T) + \int_E n(s,x,T)F(dx)\right) ds \\ C_t &= \int_0^t \mathsf{P}(s,T)^2 v(s,T)v(s,T)^{\mathrm{T}} ds \\ w(dt,dx) &= \mathsf{P}(t,T)n(t,x,T)F(dx)dt. \end{aligned}$$

Lemma 6.1 *Let* P' *be a probability measure such that* $P' \overset{loc}{\ll} P$*.*

a) $\overline{\mathsf{P}}(\cdot,T)$ *is a* P' *local martingale if and only if*

$$m(t,T) + \mathsf{P}(t,T)v(t,T)v(t,T)^T\beta_t + \int_E Y(\omega,t,x)n(t,x,T)F(dx) = r_t$$

where β*,* Y *are given by Theorem 4.2.*

b) *Let* Z *be the density process. Then if* ϕ *is a predictable process such that* $Z^c = \mathscr{E}(\int_0^\cdot \phi_s dW_s)$ *then the condition in a) is equivalent to*

$$m(t,T) + \phi_t v(t,T)^T + \int_E Y(\omega,t,x)n(t,x,T)F(dx) = r_t.$$

PROOF OF LEMMA. Since $P' \overset{\text{loc}}{\ll} P$ then P is a P' semimartingale with characteristics (B', C', w'), where

$$B' = \int_0^{\cdot} \mathsf{P}(t,T)\Big[m(t,T) + \int_E n(t,x,T)F(dx) + \mathsf{P}(t,T)v(t,T)v(t,T)^{\mathrm{T}}\beta_t \\ + \int_E (Y(\omega,t,x)-1)n(t,x,T)F(dx)\Big]dt. \tag{26}$$

The expression in brackets reduces to

$$m(t,T) + \mathsf{P}(t,T)v(t,T)v(t,T)^{\mathrm{T}}\beta_t + \int_E Y(\omega,t,x)n(t,x,T)F(dx).$$

Therefore, $\overline{\mathsf{P}}(\cdot,T)$ is a P' local martingale if the last expression is (a.s.) equal to r_t.

Since $dZ^c = Z_-\phi \cdot W$ and $d\mathsf{P}^c = \mathsf{P}_-v \cdot W$ then by the definition of β,

$$\beta = \frac{d\langle Z^c, \mathsf{P}^c\rangle}{d\langle \mathsf{P}^c, \mathsf{P}^c\rangle}\frac{1}{Z_-} = \frac{\phi v^{\mathrm{T}}}{\mathsf{P}_- v v^{\mathrm{T}}}.$$

The condition in b) now follows from a).

PROOF OF THEOREM 6.1. (Necessity) Since $P' \overset{\text{loc}}{\ll} P$, there is a predictable process ϕ and a non-negative $\tilde{\mathscr{P}}$-measurable function Y such that

$$\langle Z^c, W\rangle = (\phi Z_-)\cdot A, \text{ and } YZ_- = M_\mu^P(Z|\tilde{\mathscr{P}}) \quad P\text{- a.s.}$$

Then $\phi Z_- \in L^2_{\mathrm{loc}}(W)$ by Proposition 5.1. Let $\xi = Z - (\phi Z_-)\cdot W$. It follows that $Z = \xi + (\phi Z_-)\cdot W$ and $\langle \xi^c, W\rangle = 0$. We also have that $M_\mu^P(\Delta Z|\tilde{\mathscr{P}}) = YZ_- - Z_- = Z_-(Y-1)$. Since $(\phi Z_-)\cdot W$ is continuous,

$$\Delta\xi = (Z - (\phi Z_-)\cdot W) - (Z_- - (\phi Z_-)\cdot W) = Z - Z_- = \Delta Z$$

so that $M_\mu^P(\Delta\xi|\tilde{\mathscr{P}}) = M_\mu^P(\Delta Z|\tilde{\mathscr{P}}) = Z_-(Y-1)$. By Proposition 5.2, $Z_-(Y-1) \in G_{\mathrm{loc}}(\mu)$, and if we let $\eta = \xi - Z_-(Y-1)*(\mu-\nu)$ then $M_\mu^P(\Delta\eta|\tilde{\mathscr{P}}) = 0$. In summary, we have

$$Z = \xi + (\phi Z_-)\cdot W$$

and

$$\xi = \eta + Z_-(Y-1)*(\mu-\nu),$$

so that

$$Z = (\phi Z_-)\cdot W + Z_-(Y-1)*(\mu-\nu) + \eta \tag{27}$$

where η is a local martingale such that $M_\mu^P(\Delta\eta|\tilde{\mathscr{P}}) = 0$, and $\langle \eta^c, W\rangle = \langle \xi^c, W\rangle = 0$.

Let $R_n = \inf\left(t : Z < \frac{1}{n}\right)$ and define a process H by

$$H_t = |\phi|^2 \cdot A_t + (1-\sqrt{Y})^2 * \nu_t.$$

Since $\frac{1}{Z_-}1_{[\![0,R_n]\!]} \leq n$ by definition of R_n we have that

$$\phi Z_- \in L^2_{\mathrm{loc}}(W) \Rightarrow \phi 1_{[\![0,R_n]\!]} \in L^2_{\mathrm{loc}}(W)$$

and

$$Z_-(Y-1) \in G_{\mathrm{loc}}(\mu) \Rightarrow (Y-1)1_{[\![0,R_n]\!]} \in G_{\mathrm{loc}}(\mu).$$

Using Theorem 2.2 (note $C'(Y-1) = (1-\sqrt{Y})^2 * \nu$), conclude that $H_{R_n\wedge t} < \infty$ P-a.s. for all t. Since P', P are locally equivalent by hypothesis, $Z > 0$ P-a.s. This implies that $R_n \to \infty$ a.s. as $n \to \infty$, hence $H < \infty$ a.s.

It follows from the definition of the driving terms (W,μ) and Theorem 2.3 that the characteristics are deterministic. We can then apply Theorem 3.1 and Corollary 5.2 to conclude that all local martingales have the representation property relative to (W,μ). Since the decomposition (27) was constructed so that (19) is satisfied then Corollary 5.1 implies that $\eta_t = Z_0 (= 1)$ a.s. for all t. It follows that $\mathrm{E}^P(\mathscr{E}(L)) = 1$. Since $\overline{\mathsf{P}}$ is a P'-local martingale by hypothesis, then the Lemma implies (25).

(Sufficiency) By hypothesis, we can define a probability measure $P'_t = Z_t P_t$ for each t, where $P_t = P_{|\mathscr{F}_t}$ and $Z = \mathscr{E}(L)$. The process Z satisfies

$$\begin{aligned} Z &= 1 + (Z_-\phi)\cdot W + Z_-(Y-1) * (\mu - \nu), \\ \Delta Z &= Z_-(Y-1)\ M_\mu^P\text{-a.s.} \end{aligned}$$

Since the right hand side of the last expression is $\tilde{\mathscr{P}}$-measurable, then $M_\mu^P(\Delta Z|\tilde{\mathscr{P}}) = Z_-(Y-1)$ and hence (14) is satisfied. Denoting the P'- compensator of μ by $\nu' = \nu'(dt,dx)$, Theorem 4.1 implies that

$$\nu' = Y\nu\ \ P'\text{-a.s.}$$

Multiplying both sides by $\mathsf{P}(\cdot,T)n(\cdot,\cdot,T)$ gives $w' = Yw$, where $w' = w'(dt,dx)$ is the P'-compensator of $w(dt,dx) = \mathsf{P}(t-,T)n(t,x,T)\mu(dt,dx)$. Since $\mathsf{P}(\cdot,T)n(\cdot,\cdot,T)$ is uniformly bounded by hypothesis, we can apply Theorem 4.1 again to conclude that

$$YZ_- = M_w^P(Z|\tilde{\mathscr{P}})\ \text{a.s.}$$

We observe that (15) is satisfied, so $\mathsf{P}(\cdot,T)$ is a P'-semimartingale and its first P'-characteristic B' is given by (26). The result then follows from the Lemma above.

6.4 Application: Defaultable Bonds

In this subsection we apply our results to derive sufficient conditions for absence of arbitrage in a market for a zero coupon bond subject to default risk. We follow Schönbucher [9]. In order to model the default times and their associated loss quotas, we introduce a random measure μ_d, independent of μ, with compensator ν_d defined by

$$\nu_d(dt, dq) = q\lambda_t K(dq)dt,$$

where the intensity λ_t is continuous and finite for each t, and $K(\cdot)$ is a finite measure on $E_d = [0, 1]$.

Under the canonical setting, let the filtration $\mathbf{F}$ be generated by (W, μ, μ_d). We begin with the dynamics of the *defaultable forward rates* $\overline{f}(t, T)$, with coefficients $\overline{\alpha}$, $\overline{\sigma}$, $\overline{\delta}$ defined as in the default-free case:

$$d\overline{f}(t, T) = \overline{\alpha}(t, T)dt + \overline{\sigma}(t, T)dW_t + \int_E \overline{\delta}(t, x, T)\mu(dt, dx). \tag{28}$$

We also define the *defaultable short rate* by $\overline{r}_t = \overline{f}(t, t)$. The defaultable zero coupon bond is defined as follows:

$$\mathsf{R}(t, T) = \left(1 - \int_0^1 q\mu_d(dt, dq)\right) \exp\left(- \int_t^T \overline{f}(t, s)ds\right).$$

As in the risk-free case, we obtain the defaultable bond dynamics

$$\begin{aligned} d\mathsf{R}(t, T) = \mathsf{R}(t-, T)\Bigg[&\left(\overline{r}_t + \overline{A}(t, T) + \frac{1}{2}\|\overline{S}(t, T)\|^2\right) dt + \overline{S}(t, T)dW_t \\ &+ \int_E \left(e^{\overline{D}(t,x,T)} - 1\right) \mu(dt, dx) - \int_{[0,1]} q\mu_d(dt, dq)\Bigg], \end{aligned}$$

where

$$\begin{aligned} \overline{A}(t, T) &= - \int_t^T \overline{\alpha}(t, s)ds \\ \overline{S}(t, T) &= - \int_t^T \overline{\sigma}(t, s)ds \\ \overline{D}(t, x, T) &= - \int_t^T \overline{\delta}(t, x, s)ds. \end{aligned} \tag{29}$$

Since we assume the absence of arbitrage in the risk-free case, we obtain from the theorem the probability P' under which $\frac{\mathsf{P}(\cdot,T)}{\mathsf{B}}$ is a local martingale. The

set of P'-characteristics of $\mathsf{R}(\cdot, T)$ is

$$\begin{aligned} dB_t &= \mathsf{R}(t,T)\left(m(t,T) + \int_E Y(t,x)n(t,x,T)F(dx) - \int_{[0,1]} q\lambda_t K(dq)\right)dt \\ dC_t &= \mathsf{R}(t,T)^2\|\overline{S}(t,T)\|^2 dt \\ w(dt,dx,dq) &= \mathsf{R}(t,T)\Big(Y(t,x)n(t,x,T)F(dx) - q\lambda_t K(dq)\Big)dt, \end{aligned}$$

where

$$m(t,T) = \overline{r}_t + \overline{A}(t,T) + \frac{1}{2}\|\overline{S}(t,T)\|^2 + \phi_t\overline{S}(t,T)^T,$$

and

$$n(t,x,T) = \left(e^{\overline{D}(t,x,T)} - 1\right).$$

It follows from the proof of the theorem that for absence of arbitrage in this case we require the existence of an appropriate random variable Y_d from which we can construct an equivalent martingale measure P'' using $\mathscr{E}(L_d)$ where

$$L_d = (Y_d - 1) * (\mu_d - \nu_d). \tag{30}$$

We now seek conditions on the first P''-characteristic B'' of R so that $\frac{\mathsf{R}(\cdot,T)}{\mathsf{B}}$ is a local martingale. We have that

$$\begin{aligned} dB''_t = \mathsf{R}(t-,T)\Bigg[&\overline{r}_t + \overline{A}(t,T) + \frac{1}{2}\|\overline{S}(t,T)\|^2 + \phi_t\overline{S}(t,T)^T \\ &+ \int_E Y(t,x)\left(e^{\overline{D}(t,x,T)} - 1\right)F(dx) - \int_{[0,1]} Y_d(t,q)q\lambda_t K(dq)\Bigg]dt. \end{aligned}$$

Hence we require the following conditions:

$$\begin{aligned} \overline{A}(t,T) + \|\overline{S}(t,T)\|^2 + \phi_t\overline{S}(t,T)^T& \\ + \int_E Y(t,x)\left(e^{\overline{D}(t,x,T)} - 1\right)F(dx) &= 0 \end{aligned} \tag{31}$$

and

$$0 < \overline{r}_t - r_t = \int_{[0,1]} Y_d(t,q)q\lambda_t K(dq). \tag{32}$$

The last inequality is a formal relationship between between the default-free and defaultable models necessary for absence of arbitrage. The following result summarizes the above discussion.

Theorem 6.3 *Let the defaultable forward rate dynamics be given by (28). There exists a martingale measure if the following conditions hold:*

(i) The conditions in Theorem 6.2 hold for the default-free forward rates.

(ii) There exists a predictable function $Y_d(t,q)$ with $Y_d > 0$ satisfying

$$\int_0^t \int_{[0,1]} |Y(s,q)| K(dq) ds < \infty, \quad \textit{and} \quad E^{P'}(\mathscr{E}(L_d)_t) = 1$$

for all finite t. Here L_d is given by (30).

(iii) The coefficients $\overline{A}(t,T)$, $\overline{S}(t,T)$ and $\overline{D}(t,x,T)$ satisfy (31) and (32).

7 Concluding Remarks

We have described a mathematical framework in which to study term structure models. All processes considered are semimartingales, and we represent them in terms of their set of characteristics. In particular, the necessary and sufficient conditions for absence of arbitrage in a bond market are presented in terms of the characteristics of the price process. As an example, we presented a stable process as the underlying source of randomness in terms of its characteristics. We also illustrate how the general methodology can be extended to defaultable bonds. In future work we will seek the development of estimation and numerical procedures following this approach.

References

[1] S. Babbs, N. Webber, "A Theory of the Term Structure with an official Short Rate", preprint, University of Warwick, (1994).

[2] T. Björk, Y. Kabanov, W. Runggaldier, "Bond Market Structure in the Presence of Marked Point Processes", *Math. Finance* 7, 211-239, (1997).

[3] L. El-Jahel, H. Lindberg, W. Perraudin, "Yield Curves with Jump Short Rates", preprint, (2001).

[4] D. Heath, R. Jarrow, A. Morton, "Bond Pricing and the Term Structure of Interest Rates", *Econometrica* 60(1), 77-106, (1992).

[5] S.R. Hurst, E. Platen, S.T. Rachev, "Option Pricing for a Logstable Asset Price Model", *Mathematical and Computer Modelling* 29, 105-119, (1999).

[6] R. Jarrow, D. Madan, "Option Pricing Using the Term Structure of Interest Rates to Hedge Systematic Discontinuities in Asset Returns", *Math. Finance* 5, 311-336, (1995).

[7] G. Samorodnitsky, M.S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York, (1994).

[8] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, (2000).

[9] P.J. Schönbucher, *Credit Risk Modelling and Credit Derivatives*, Ph.D. dissertation, University of Bonn, (2000).

[10] H. Shirakawa, "Interest Rate Option Pricing with Poisson-Gaussian Forward Rate Curve Processes", *Math. Finance* 1, 77-94, (1991).

[11] A.N. Shiryaev, J. Jacod, *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin, (1987).

[12] A.N. Shiryaev, *Essentials of Stochastic Finance: Facts, Models, Theory*, World Scientific, Singapore, (1999).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.4,369-396,2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Univariate stable laws in the field of finance — parameter estimation

Stoyan Stoyanov
Chief Financial Researcher, FinAnalytica Inc.
stoyan.stoyanov@finanalytica.com

Borjana Racheva-Iotova
MD European Operations, FinAnalytica Inc.
borjana.racheva@finanalytica.com

Abstract

The seminal work of Mandelbrot and Fama, carried out in the sixties, suggested the class of α-stable laws as a probabilistic model of financial assets returns. Stable distributions possess several properties which make plausible their application in the field of finance — heavy tails, excess kurtosis, domains of attraction. The problem of parameter estimation arises naturally and was first tackled by Mandelbrot and Fama. In this paper we summarize some of the basic approaches and provide comparison between the estimators using bootstrap in empirical examples.

Keywords stable distributions, finance, asset returns, parameter estimation

2000 AMS Subject Classification 60F05, 91B28, 62F10

1 Introduction

Stable laws are a family of distributions which arises from generalizations of the central limit theorem. Stable non-Gaussian distributions possess several properties that make them attractive in the applications, namely heavy tails, excess kurtosis and domains of attraction. Particularly in the field of finance, the class of stable laws is attractive because it is empirically observed that assets returns are heavy-tailed and leptokurtic. The stable Paretian hypothesis, namely the conjecture that financial assets returns have stable non-Gaussian distributions, was first suggested by Mandelbrot and Fama in the 60's. Subsequent studies confirmed the results of the initial research which supported the stable Paretian hypothesis.

A well-known property of stable non-Gaussian distributions is that they do not possess a finite second moment. Certainly the application of infinite-variance distributions as theoretical models of bounded variables such as financial assets returns seems inappropriate. Moreover any empirical distribution has a finite variance, hence it may seem that infinite variance distributions are inapplicable in any context. Nevertheless there is ample empirical evidence that the probability of large deviations of the changes in stock market prices is so great that any statistical theory based on finite variance distributions is impossible to predict accurately. As it is remarked in [9] and [4], the sum of a large number of these variables is often dominated by one of the summands which is a theoretical property of infinite variance distributions. Hence an infinite variance distribution may be an appropriate probabilistic model of such variables.

A random variable X is said to have stable distribution if for any $n \geq 2$, there is a positive number C_n and a real number D_n such that the sum of n independent copies of X, $X_1 + X_2 + \ldots + X_n$, has the same distribution as $C_n X + D_n$. Equivalently, a random variable X is said to have stable distribution, which is denoted by $X \sim S_\alpha(\sigma, \beta, \mu)$, if its characteristic function (ch.f.) admits the representation P_0:

$$\varphi(t) = Ee^{itX} = \begin{cases} \exp\{-\sigma^\alpha |t|^\alpha (1 - i\beta \frac{t}{|t|} \tan(\frac{\pi\alpha}{2})) + i\mu t\}, & \alpha \neq 1 \\ \exp\{-\sigma |t| (1 + i\beta \frac{2}{\pi} \frac{t}{|t|} \ln(|t|)) + i\mu t\}, & \alpha = 1 \end{cases} \tag{1}$$

where $0 < \alpha \leq 2$ is the index of stability, $-1 \leq \beta \leq 1$ is the skewness parameter, $\sigma > 0$ is a scale parameter and $\mu \in \mathbb{R}$ is a location parameter. For a discussion of the properties of the parameters of stable laws and the properties of the class of α-stable distributions in financial context, the reader is referred to [14]. An excellent reference of the properties of univariate and multivariate stable laws is [13].

Using the properties of stable laws, it is straightforward to establish the following useful expressions:

$$f(x; \alpha, \beta) = f(-x; \alpha, -\beta), \quad F(x; \alpha, \beta) = 1 - F(-x; \alpha, -\beta) \tag{2}$$

and

$$f(x) = \frac{1}{\sigma} f\Big(\frac{x - \mu}{\sigma}; \alpha, \beta\Big), \quad F(x) = F\Big(\frac{x - \mu}{\sigma}; \alpha, \beta\Big) \tag{3}$$

where $f(x; \alpha, \beta)$ and $F(x; \alpha, \beta)$ are the probability density function (p.d.f.) and the cumulative distribution function (c.d.f.) of $X \sim S_\alpha(\sigma, \beta, \mu)$ respectively. For more details the reader is referred to [14].

Among the class of infinite-variance distributions, stable non-Gaussian laws are very attractive because they have domains of attraction, i.e. sums of properly normalized independent identically distributed random variables converge in distribution to a stable law. This fact has important implications in financial modeling, for a discussion see [12] and [14].

The problem of parameter estimation of stable distributions was first tackled by Mandelbrot then by Fama and Roll, [5] and [6]. This is a non-trivial task because, with a few exceptions, there are no closed-form expressions for the density and distribution functions. For example the classical maximum likelihood method in this case depends on numerical approximations of the density and could be extremely time-consuming, for a discussion see [14]. Moreover standard estimation techniques based on asymptotic results which rely on a finite second moment are irrelevant. In this paper we consider some approaches to the estimation of stable parameters, namely quantile methods, ch.f. based methods and maximum likelihood. Then we use empirical data to compare the performance of the estimators in terms of standard deviation of the estimates and box-whiskers diagrams. This is achieved by means of resampling taking advantage of the non-parametric bootstrap method.

The parametrization is an important issue as the classical representation of the ch.f. in equation (1) is not continuous in the parameter space. In numerical and statistical work, it is recommended to work with a continuous parametrization, for this reason, whenever appropriate, we shall use the following representation P_1 of the ch.f.

$$\varphi(t) = \begin{cases} \exp\{-|\sigma t|^\alpha + i\sigma t\beta(|\sigma t|^{\alpha-1} - 1)\tan(\frac{\pi\alpha}{2}) + i\mu_1 t\}, & \alpha \neq 1 \\ \exp\{-|\sigma t| + i\sigma t\beta\frac{2}{\pi}\ln|\sigma t| + i\mu_1 t\}, & \alpha = 1 \end{cases} \tag{4}$$

The parameters in equation (1) and (4) are related by

$$\mu_1 = \begin{cases} \mu + \beta\sigma\tan\frac{\pi\alpha}{2}, & \alpha \neq 1 \\ \mu, & \alpha = 1 \end{cases}$$

The relationships between these and other parametrizations are discussed in [14] and [15].

2 Parameter estimation

Generally speaking, parameter estimation techniques for the class of stable laws come into three categories — quantile methods, characteristic function

based methods and maximum likelihood. The approaches from the first type use predetermined empirical quantiles to estimate stable parameters. For example the method of Fama and Roll [5], [6] for symmetric α-stable distributions and its modified version of McCulloch [10] for the skewed case belong to this group.

Ch.f. based methods include the method of moments approach suggested by Press [11] and regression-type procedures proposed by Koutrouvelis [8] and Kogon and Williams [1]. Simulation studies available in the literature ([1], [2]), show the superiority of the regression-type estimation over the quantile methods.

The validity of maximum likelihood estimation (MLE) theory was demonstrated by DuMouchel [3]. The comparison studies between MLE and the quantile method of McCulloch in [12] recommend the maximum likelihood estimator.

In this paper we shall review and compare McCulloch's quantile method, the method of moments, the regression-type estimator of Kogon and Williams and MLE.

2.1 Quantile method of McCulloch

The estimation procedure proposed by McCulloch in [10] is a generalization of the quantile method of Fama and Roll in [5], [6] for the symmetric case. The estimates of stable parameters in parametrization P_0 are consistent and asymptotically normal if $0.6 < \alpha \leq 2$. We shall adopt the standard notation for theoretical and empirical quantiles, namely x_p is the p-th quantile if $F(x_p) = p$, where $F(x)$ is the c.d.f. of a random variable and given a sample of observations $x_1, x_2, \ldots, x_n$, then $\hat{x}_p$ is the sample quantile if $F_n(\hat{x}_p) = p$, where $F_n(x)$ is the sample c.d.f..

According to [10], let us define two functions of theoretical quantiles :

$$v_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}}$$

$$v_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.50}}{x_{0.95} - x_{0.05}}$$

The functions v_α and v_β have this special form because by expression (3) it appears that they do not depend on the scale and the location parameter, i. e.

$$\left| \begin{array}{l} v_\alpha = \phi_1(\alpha, \beta) \\ v_\beta = \phi_2(\alpha, \beta) \end{array} \right. \tag{5}$$

		v_β						
		0	0.1	0.2	0.3	0.5	0.7	1
v_α	2.439	2	2	2	2	2	2	2
	2.5	1.916	1.924	1.924	1.924	1.924	1.924	1.924
	2.6	1.808	1.813	1.829	1.829	1.829	1.829	1.829
	2.7	1.729	1.73	1.737	1.745	1.745	1.745	1.745
	2.8	1.664	1.663	1.663	1.668	1.676	1.676	1.676
	3	1.563	1.56	1.553	1.548	1.547	1.547	1.547
	3.2	1.484	1.48	1.471	1.46	1.448	1.438	1.438
	3.5	1.391	1.386	1.378	1.364	1.337	1.318	1.318
	4	1.279	1.273	1.266	1.25	1.21	1.184	1.15
	5	1.128	1.121	1.114	1.101	1.067	1.027	0.973
	6	1.029	1.021	1.014	1.004	0.974	0.935	0.874
	8	0.896	0.892	0.887	0.883	0.855	0.823	0.769
	10	0.818	0.812	0.806	0.801	0.78	0.756	0.691
	15	0.698	0.695	0.692	0.689	0.676	0.656	0.595
	25	0.593	0.59	0.588	0.586	0.579	0.563	0.513

Table I: $\alpha = \psi_1(v_\alpha, v_\beta) = \psi_1(v_\alpha, -v_\beta)$

Employing equation (2), we have that $F(-x_p; \alpha, -\beta) = F(x_{1-p}; \alpha, \beta)$ and therefore we have the relations:

$$\begin{aligned} \phi_1(\alpha, \beta) &= \phi_1(\alpha, -\beta) \\ \phi_2(\alpha, \beta) &= -\phi_2(\alpha, -\beta) \end{aligned} \tag{6}$$

The system of equations (5) can be inverted and the parameters α and β can be expressed as functions of the quantities v_α and v_β:

$$\left| \begin{array}{l} \alpha = \psi_1(v_\alpha, v_\beta) \\ \beta = \psi_2(v_\alpha, v_\beta) \end{array} \right. \tag{7}$$

Replacing v_α and v_β in equations (7) with their sample counterparts $\hat{v}_\alpha$ and $\hat{v}_\beta$:

$$\hat{v}_\alpha = \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}}$$

$$\hat{v}_\beta = \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.50}}{\hat{x}_{0.95} - \hat{x}_{0.05}}$$

yields estimators $\hat{\alpha}$ and $\hat{\beta}$:

		v_β						
		0	0.1	0.2	0.3	0.5	0.7	1
v_α	2.439	0	2.16	1	1	1	1	1
	2.5	0	1.592	3.39	1	1	1	1
	2.6	0	0.759	1.8	1	1	1	1
	2.7	0	0.482	1.048	1.694	1	1	1
	2.8	0	0.36	0.76	1.232	2.229	1	1
	3	0	0.253	0.518	0.823	1.575	1	1
	3.2	0	0.203	0.41	0.632	1.244	1.906	1
	3.5	0	0.165	0.332	0.499	0.943	1.56	1
	4	0	0.136	0.271	0.404	0.689	1.23	2.195
	5	0	0.109	0.216	0.323	0.539	0.827	1.917
	6	0	0.096	0.19	0.284	0.472	0.693	1.759
	8	0	0.082	0.163	0.243	0.412	0.601	1.596
	10	0	0.074	0.174	0.22	0.377	0.546	1.482
	15	0	0.064	0.128	0.191	0.33	0.478	1.362
	25	0	0.056	0.112	0.167	0.285	0.428	1.274

Table II: $\beta = \psi_2(v_\alpha, v_\beta) = -\psi_2(v_\alpha, -v_\beta)$

$$\left| \begin{array}{l} \hat{\alpha} = \psi_1(\hat{v}_\alpha, \hat{v}_\beta) \\ \hat{\beta} = \psi_2(\hat{v}_\alpha, \hat{v}_\beta) \end{array} \right. \tag{8}$$

The functions $\psi_1(.)$ and $\psi_2(.)$ are tabulated in Tables I and II. It should be noted that because of property (6), we have that:

$$\psi_1(v_\alpha, v_\beta) = \psi_1(v_\alpha, -v_\beta)$$
$$\psi_2(v_\alpha, v_\beta) = -\psi_2(v_\alpha, -v_\beta)$$

In other words, the sign of $\hat{v}_\beta$ determines the sign of β.

Since there are no closed-form expressions for the functions $\psi_1(.)$ and $\psi_2(.)$, we compute estimates of α and β from the statistics $\hat{v}_\alpha$ and $\hat{v}_\beta$ using the values in Tables I and II and linear interpolation for intermediate values. If it happens that $\hat{v}_\alpha$ is below 2.439, $\hat{\alpha}$ should be set equal to 2 and $\hat{\beta}$ equal to zero. Table II contains values larger than 1 for more precise interpolation. If $\hat{\beta} > 1$, it should be reduced to 1.

McCulloch provides estimator for the scale parameter σ which is very similar to the estimator given by Fama and Roll. Let us first define v_σ as:

$$v_\sigma = \frac{x_{0.75} - x_{0.25}}{\sigma} = \phi_3(\alpha, \beta)$$

		β				
		0	0.25	0.5	0.75	1
α	0.5	2.588	3.073	4.534	6.636	9.144
	0.6	2.337	2.635	3.542	4.808	6.247
	0.7	2.189	2.392	3.004	3.844	4.775
	0.8	2.098	2.244	2.676	3.265	3.912
	0.9	2.04	2.149	2.461	2.886	3.356
	1	2	2.085	2.311	2.624	2.973
	1.1	1.98	2.04	2.205	2.435	2.696
	1.2	1.965	2.007	2.125	2.294	2.491
	1.3	1.955	1.984	2.067	2.188	2.333
	1.4	1.946	1.967	2.022	2.106	2.211
	1.5	1.939	1.952	1.988	2.045	2.116
	1.6	1.933	1.94	1.962	1.997	2.043
	1.7	1.927	1.93	1.943	1.961	1.987
	1.8	1.921	1.922	1.927	1.936	1.947
	1.9	1.914	1.915	1.916	1.918	1.921
	2	1.908	1.908	1.908	1.908	1.908

Table III: $v_\sigma = \phi_3(\alpha,\beta) = \phi_3(\alpha,-\beta)$

The function $\psi_3(\alpha,\beta)$ is given in Table III. Employing the same arguments that led us to equations (6) yields the relation $\phi_3(\alpha,\beta) = \phi_3(\alpha,-\beta)$. The estimator $\hat{\sigma}$ is received after replacing α and β with the estimates found according to equations (8):

$$\hat{\sigma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\phi_3(\hat{\alpha},\hat{\beta})}$$

Estimation of the location parameter μ is a more involved affair because of the discontinuity of the parametric representation of the ch.f. P_0 when $\alpha \to 1$ and $\beta \neq 0$. First we estimate the shifted location parameter ζ defined by:

$$\zeta = x_{0.50} + \sigma sign(\beta)\phi_4(\alpha,\beta) \tag{9}$$

where $\phi_4(\alpha,\beta)$ is tabulated in Table IV and has the property $\phi_4(\alpha,\beta) = \phi_4(\alpha,-\beta)$. The location parameter μ is related to ζ according to:

$$\mu = \begin{cases} \zeta - \beta\sigma \tan\frac{\pi\alpha}{2}, & \alpha \neq 1 \\ \zeta, & \alpha = 1 \end{cases} \tag{10}$$

Replacing the parameters in equations (9) and (10) with their sample counterparts yields the estimator $\hat{\mu}$:

		β				
		0	0.25	0.5	0.75	1
	0.5	0	-0.061	-0.279	-0.659	-1.198
	0.6	0	-0.078	-0.272	-0.581	-0.997
	0.7	0	-0.089	-0.262	-0.52	-0.853
	0.8	0	-0.096	-0.25	-0.469	-0.742
	0.9	0	-0.099	-0.237	-0.424	-0.652
	1	0	-0.098	-0.223	-0.383	-0.576
	1.1	0	-0.095	-0.208	-0.346	-0.508
	1.2	0	-0.09	-0.192	-0.31	-0.447
α	1.3	0	-0.084	-0.173	-0.276	-0.39
	1.4	0	-0.075	-0.154	-0.241	-0.335
	1.5	0	-0.066	-0.134	-0.206	-0.283
	1.6	0	-0.056	-0.111	-0.17	-0.232
	1.7	0	-0.043	-0.088	-0.132	-0.179
	1.8	0	-0.03	-0.061	-0.092	-0.123
	1.9	0	-0.017	-0.032	-0.049	-0.064
	2	0	0	0	0	0

Table IV: $\phi_4(\alpha,\beta) = \phi_4(\alpha,-\beta)$

$$\hat{\zeta} = \hat{x}_{0.50} + \hat{\sigma} sign(\hat{\beta})\phi_4(\hat{\alpha},\hat{\beta})$$

and

$$\hat{\mu} = \begin{cases} \hat{\zeta} - \hat{\beta}\hat{\sigma}\tan\frac{\pi\hat{\alpha}}{2}, & \hat{\alpha} \neq 1 \\ \hat{\zeta}, & \hat{\alpha} = 1 \end{cases}$$

It should be observed that a significant advantage of the method considered is the lack of heavy computations. On the personal homepage of McCulloch (http://www.econ.ohio-state.edu/jhm/jhm.html) a FORTRAN implementation of the algorithm is publicly available.

2.2 Ch.f. based methods

The characteristic function based methods rely on the sample ch.f. for parameter estimation. The sample ch.f. is defined as:

$$\hat{\varphi}(t) = \frac{1}{n}\sum_{j=1}^{n} e^{itx_j}, \quad t \in \mathbb{R} \tag{11}$$

where $x_1, x_2, \ldots, x_n$ is a sample of independent, identically distributed (iid) observations on a random variable X. Since $|\hat{\varphi}(t)| \leq 1$, all moments of the

random variable $\hat{\varphi}(t)$ are finite and, according to equation (11), for any t it is the sample mean of the iid random variables e^{itx_j}. As a consequence, from the law of large numbers, it can be inferred that the sample ch.f. is a consistent estimator of the ch.f. $\varphi_X(t) = \mathbb{E}e^{itX}$, $t \in \mathbb{R}$ of a random variable X.

2.2.1 The method of moments

Press [11] suggested a simple and straightforward approach to estimation of parameters of stable laws which was called the method of moments. His approach is based on certain transformations of the ch.f. in parametrization P_0. From the parametric representation (1) it follows that

$$|\varphi(t)| = \exp(-\sigma^\alpha |t|^\alpha), \quad t \in \mathbb{R} \tag{12}$$

and therefore $-\ln|\varphi(t)| = \sigma^\alpha |t|^\alpha$ for any real t.

Case $\alpha \neq 1$. If we choose t_1 and t_2 such that $t_1 \neq t_2 \neq 0$, we have the following system of two equations:

$$\left| \begin{array}{l} -\ln|\varphi(t_1)| = \sigma^\alpha |t_1|^\alpha \\ -\ln|\varphi(t_2)| = \sigma^\alpha |t_2|^\alpha \end{array} \right.$$

which can be solved for α and σ. Replacing the ch.f. for its sample equivalent $\hat{\varphi}(t)$ yields the estimators $\hat{\alpha}$ and $\hat{\sigma}$:

$$\hat{\alpha} = \frac{\ln \frac{\ln|\hat{\varphi}(t_1)|}{\ln|\hat{\varphi}(t_2)|}}{\ln|\frac{t_1}{t_2}|} \tag{13}$$

and

$$\ln\hat{\sigma} = \frac{\ln|t_1| \ln(-\ln|\hat{\varphi}(t_2)|) - \ln|t_2| \ln(-\ln|\hat{\varphi}(t_1)|)}{\ln\left|\frac{\hat{\varphi}(t_1)}{\hat{\varphi}(t_2)}\right|} \tag{14}$$

Estimation of the skewness and the location parameter requires more efforts. Let us first denote the imaginary part of the logarithm of the ch.f. in P_0 as $u(t)$:

$$u(t) = \Im(\ln\varphi(t)) = \mu t + \sigma^\alpha |t|^\alpha \beta sign(t) \tan\frac{\pi\alpha}{2}$$

Then if we choose two non-zero values t_3 and t_4 such that $t_3 \neq t_4$ we can write a system of two equations:

$$\left| \begin{array}{l} \frac{u(t_3)}{t_3} = \mu + \sigma^\alpha |t_3|^{\alpha-1} \beta \tan \frac{\pi\alpha}{2} \\ \frac{u(t_4)}{t_4} = \mu + \sigma^\alpha |t_4|^{\alpha-1} \beta \tan \frac{\pi\alpha}{2} \end{array} \right.$$

It is possible to solve the system for β and μ and again replacing α, σ and $u(t)$ with their sample counterparts yields the required estimators. Since

$$\hat{\phi}_\xi(t) = \left(\frac{1}{n} \sum_{j=1}^{n} \cos tx_j \right) + i \left(\frac{1}{n} \sum_{j=1}^{n} \sin tx_j \right)$$

and taking advantage of the properties of complex numbers we achieve the estimator $\hat{u}(t)$:

$$\tan \hat{u}(t) = \frac{\sum_{j=1}^{n} \sin tx_j}{\sum_{j=1}^{n} \cos tx_j}$$

Finally for $\hat{\beta}$ and $\hat{\mu}$ we have:

$$\hat{\beta} = \frac{\frac{\hat{u}(t_4)}{t_4} - \frac{\hat{u}(t_3)}{t_3}}{\left[|t_4|^{\hat{\alpha}-1} - |t_3|^{\hat{\alpha}-1} \right] \hat{\sigma}^{\hat{\alpha}} \tan \frac{\pi\hat{\alpha}}{2}} \tag{15}$$

and

$$\hat{\mu} = \frac{|t_4|^{\hat{\alpha}-1} \frac{\hat{u}(t_3)}{t_3} - |t_3|^{\hat{\alpha}-1} \frac{\hat{u}(t_4)}{t_4}}{|t_4|^{\hat{\alpha}-1} - |t_3|^{\hat{\alpha}-1}} \tag{16}$$

Case $\alpha = 1$. If $\alpha = 1$, equation (12) allows us to construct the estimator $\hat{\sigma}$ directly:

$$\hat{\sigma} = -\frac{\ln |\varphi(t_1)|}{t_1}$$

where $t_1 \neq 0$. Similar arguments as in the case $\alpha \neq 1$ lead us to:

$$\hat{\beta} = \frac{\frac{\hat{u}(t_3)}{t_3} - \frac{\hat{u}(t_4)}{t_4}}{\frac{2}{\pi} \hat{\sigma} \ln |\frac{t_4}{t_3}|}$$

$$\hat{\mu} = \frac{\ln |t_4| \frac{\hat{u}(t_3)}{t_3} - \ln |t_3| \frac{\hat{u}(t_4)}{t_4}}{\ln |t_4| - \ln |t_3|}$$

where $t_3 \neq t_4$ and both are non-zero.

The estimators of stable parameters are consistent since they are based on $\hat{\varphi}(t)$, $\Re\hat{\varphi}(t)$ and $\Im\hat{\varphi}(t)$ which are consistent estimators of $\varphi(t)$, $\Re\varphi(t)$ and

$\Im\varphi(t)$ by the law of large numbers. The question which still remains is the best way to choose $t_1, \ldots, t_4$, since obviously the derived estimators are not invariant of their choice. Koutrouvelis in his simulation studies in [7] uses the values $t_1 = 0.2$, $t_2 = 0.8$, $t_3 = 0.1$ and $t_4 = 0.4$, which are selected for the normalized case ($\sigma = 1, \mu = 0$). Because of the following property of the ch.f. of an arbitrary random variable X:

$$\varphi_{\sigma X+\mu}(t) = e^{it\mu}\varphi_X(\sigma t)$$

it is clear that for different σ and μ we shall have to choose different values for $t_1, \ldots, t_4$ to achieve equal performance, i.e. the values determined for the normalized case will not be equally "good" for a non-normalized case. For this reason, if we aim at estimation of stable parameters by the method of moments, we need first to find initial estimates of the scale and the location parameter and to normalize the sample. Without incurring significant additional computational burden, initial estimates could be computed with the help of a quantile method. For such purposes Koutrouvelis [7] uses the method of Fama and Roll [5], [6] despite the bias in the estimate of σ even in the symmetric case. We shall adopt his approach in our computations.

To summarize, the algorithm for estimating stable parameters by the method of moments, given a sample of iid observations $x_1, x_2, \ldots, x_n$, is as follows:

1. Compute initial estimates $\hat{\sigma}_0$ and $\hat{\mu}_0$ of σ and μ respectively, according to:

$$\hat{\sigma}_0 = \frac{\hat{x}_{0.72} - \hat{x}_{0.28}}{1.654}$$

   and $\hat{\mu}_0$ equals the 50% truncated sample average - the mean of the middle 50% of the ordered observations.

2. Normalize the sample with the initial estimates:

$$x'_k = (x_k - \hat{\mu}_0)/\hat{\sigma}_0, \quad k = 1, 2, \ldots n$$

3. Using the normalized sample $x'_1, x'_2, \ldots, x'_n$, calculate $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}_1$ and $\hat{\mu}_1$ according to equations (13), (15), (14) and (16) respectively.

4. Compute the final estimates $\hat{\sigma}$ and $\hat{\mu}$:

$$\hat{\sigma} = \hat{\sigma}_0\hat{\sigma}_1, \qquad \hat{\mu} = \hat{\sigma}_0\hat{\mu}_1 + \hat{\mu}_0$$

2.2.2 The regression-type estimator of Kogon-Williams

Regression-type estimators are also based on the sample ch.f.. It is possible to derive simple expressions, linear with respect to stable parameters, and construct estimators using the least squares technique. Kogon and Williams suggest such a procedure in [1] with the ch.f. being parametrized according to the continuous parametrization P_1 defined in equation (4). Their approach is similar to the that of Koutrouvelis in [7] and [8].

The linear equations follow directly from the convenient form of the logarithm of the ch.f.:

$$\ln[-\Re(\ln\varphi(t))] = \alpha\ln\sigma + \alpha\ln|t| \tag{17}$$

$$\Im(\ln\varphi(t)) = \mu_1 t + \beta\sigma t(|\sigma t|^{\alpha-1} - 1)\tan\frac{\pi\alpha}{2} \tag{18}$$

Estimators of the stable parameters can be constructed using the method of least squares after replacing the ch.f. for the sample ch.f.. Certainly here we face the same problem as in the method of moments - the sample ch.f. should be evaluated for certain values of the argument. Koutrouvelis gives tables in [7] and [8] which relate the values of the sample ch.f. argument to the value of the index of stability α and the sample size. The major advantage of the procedure in [1] is that the provided values of the sample ch.f. argument are invariant of any other parameters. Having conducted numerous experiments, Kogon and Williams report in [1] that the most suitable choice is $t_k = \{0.1 + 0.1k, \quad k = 0, 1, \ldots, 9\}$ - 10 equally spaced points in the interval $[0.1, 1]$. Undoubtedly the sample should be normalized before applying the method of least squares, otherwise the optimal selection of the sample ch.f. arguments would depend on the scale and the modified location parameter. For preliminary estimation of σ and μ, it is suggested to use the quantile method of McCulloch.

The algorithm is as follows:

1. Given a sample of iid observations $x_1, x_2, \ldots, x_n$ first we find preliminary estimates σ_0 and μ_{01} utilizing the quantile method of McCulloch and we normalize the observations:

$$x'_j = \frac{x_j - \hat{\mu}_{01}}{\hat{\sigma}_0}, \quad j = 1, 2, \ldots n$$

2. Next we consider the regression equation constructed from equation (17):

$$y_k = b + \alpha w_k + \epsilon_k, \quad k = 0, 1, \ldots 9$$

where $y_k = \ln[-\Re(\ln \hat{\varphi}(t_k))]$, $w_k = \ln|t_k|$, $t_k = \{0.1 + 0.1k, \quad k = 0, 1, \ldots, 9\}$ and ϵ_k denotes the error term. We find $\hat{\alpha}$ and $\hat{b}$ according to the method of least squares using the normalized sample $x'_1, x'_2, \ldots, x'_n$. The estimator $\hat{\sigma}_1$ of the scale parameter of the normalized sample is:

$$\hat{\sigma}_1 = \exp\left(\frac{\hat{b}}{\hat{\alpha}}\right)$$

3. Estimators $\hat{\beta}$ and $\hat{\mu}_{11}$ of the skewness parameter and the modified location parameter respectively are derived from the second regression equation based on (18):

$$z_k = \mu_{11} t_k + \beta v_k + \eta_k, \quad k = 0, 1, \ldots 9$$

where $z_k = \Im(\ln \hat{\varphi}(t_k))$, $v_k = \hat{\sigma}_1 t_k(|\hat{\sigma}_1 t_k|^{\hat{\alpha}-1} - 1) \tan \frac{\pi\hat{\alpha}}{2}$, $t_k = \{0.1 + 0.1k, \quad k = 0, 1, \ldots, 9\}$ and η_k is the error term.

4. The final estimators $\hat{\sigma}$ and $\hat{\mu}_1$ proceed from:

$$\hat{\sigma} = \hat{\sigma}_0 \hat{\sigma}_1, \quad \hat{\mu}_1 = \hat{\mu}_{01} + \hat{\sigma}_0 \hat{\mu}_{11}$$

If we aim at estimating the location parameter μ, we need to take advantage of the connection between the two parametric forms P_0 and P_1:

$$\hat{\mu} = \hat{\mu}_1 - \hat{\beta}\hat{\sigma} \tan \frac{\pi\hat{\alpha}}{2}$$

In [1] there is a huge Monte Carlo study in which the method of Kogon-Williams is compared to the approach of Koutrouvelis [7], [8]. The result is that from computational viewpoint the former is more efficient. It is definitely superior to the latter when α is close to zero and $\beta \neq 0$. The approach of Koutrouvelis outperforms that of Kogon-Williams only in the estimation of β.

2.3 Maximum likelihood

The method of maximum likelihood is very attractive because of the good asymptotic properties of the estimates, provided that the likelihood function obeys certain general conditions. The likelihood function is defined as:

$$L(x_1, x_2, \ldots x_n|\theta) = \prod_{k=1}^{n} f(x_k|\theta)$$

where $x_1, x_2, \ldots x_n$ is a sample of iid observations of a random variable X, $f(x|\theta)$ is the p.d.f. of X, and θ is a vector of parameters. In the case of stable distributions, $\theta = (\alpha, \beta, \sigma, \mu)$. Maximum likelihood estimates are found by searching for that parameter values which maximize the likelihood function, or equivalently, the log-likelihood function:

$$\hat{\theta}_n = \arg\max_{\theta} \log(L(x_1, x_2, \ldots x_n|\theta)) \tag{19}$$

DuMouchel studied the applicability of maximum likelihood theory in the case of α-stable distributions in [3] by verifying whether the likelihood function complies with a set of conditions that guarantee the validity of the theory. The theorem proved in the paper and adapted to parametrization P_0 is the following

Theorem 1. *When sampling from a stable distribution, $\hat{\theta}_n$, the maximum likelihood estimate for $\theta = (\alpha, \beta, \sigma, \mu)$ based on the first n observations, restricted so that $\hat{\alpha}_n$, the estimate of α, satisfies $\hat{\alpha}_n > \epsilon$, ϵ arbitrarily small and positive, is consistent and asymptotically normal as long as θ_0, the true value of θ is in the interior of the parameter space (that is the cases $\alpha_0 \leq \epsilon$, $\alpha_0 = 2$ and $|\beta| = 1$ are excluded) and the additional case $(\alpha_0 = 1, \beta_0 \neq 0)$ is excluded.*

Clearly if we intend to derive expressions for MLE analytically, we need to have closed-form expressions for the p.d.f.s of stable laws. Such expressions are not known to exist in the general case and the problem of MLE of stable parameters should be attacked numerically, i.e. we have to numerically search for the solution of problem (19) in which the p.d.f. is approximated. Approaches to numerical approximation of the p.d.f.s and the c.d.f.s of stable laws are considered in [14]. In the empirical part of this paper, we use the FFT-method combined with a Bergström series expansion for tail approximation.

3 Performance comparison

The performance of the estimators discussed in the previous section is investigated in an empirical study. We have selected two equities - Microsoft and DVI Inc., one index - Dow Jones Industrial Average and the JPY/USD exchange rate. The selection criteria was to have certain diversity in the parameter space.

The estimators are compared in terms of boxplot diagrams computed via bootstrap through non-parametric resampling. In addition to the boxplot diagrams, we provide the mean and the standard deviation of the estimators. As a goodness of fit measure, we compute the Kolmogorov distance defined by:

$$KD = \sup_x |F(x) - F_n(x)| \tag{20}$$

where $F_n(x)$ is the sample c.d.f. and $F(x)$ is the c.d.f. of the fitted distribution. Since we are dealing with heavy-tailed distributions, we also report the Anderson-Darling statistics which puts more weight on the tails of the distribution:

$$ADS = \sup_x \frac{|F(x) - F_n(x)|}{\sqrt{F(x)(1 - F(x))}} \tag{21}$$

Finally, in order to verify the superiority of stable modeling, we compare the goodness of fit when the fitted distribution is Gaussian or α-stable.

In the four empirical examples, the samples include the daily log-returns from 01.01.1999 to 31.12.2002 totaling 1042 observations. All boxplot diagrams are based on 500 resampled estimates of stable parameters.

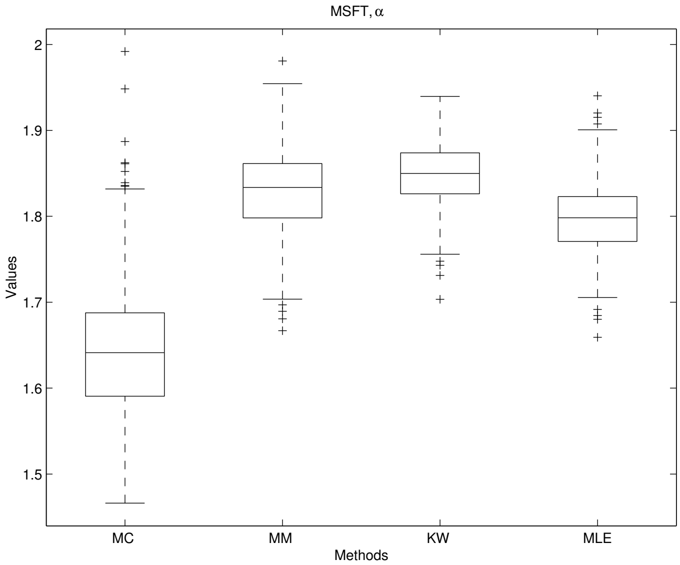


Figure 1: Boxplot diagrams of $\hat{\alpha}$, Microsoft

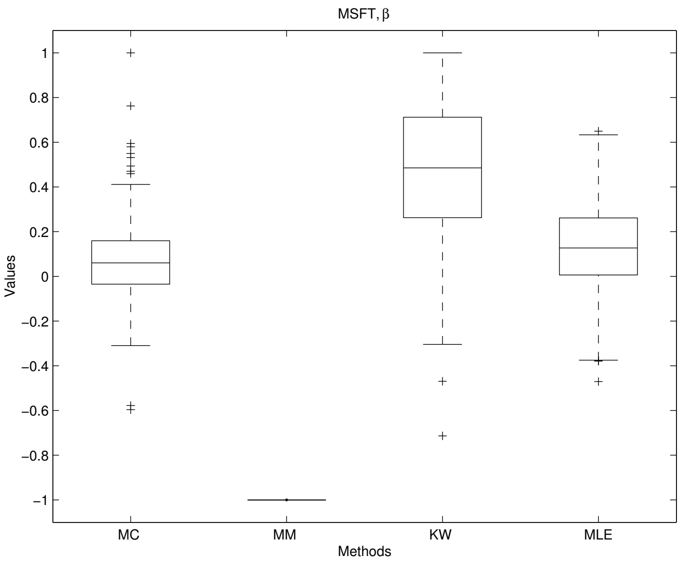


Figure 2: Boxplot diagrams of $\hat{\beta}$, Microsoft

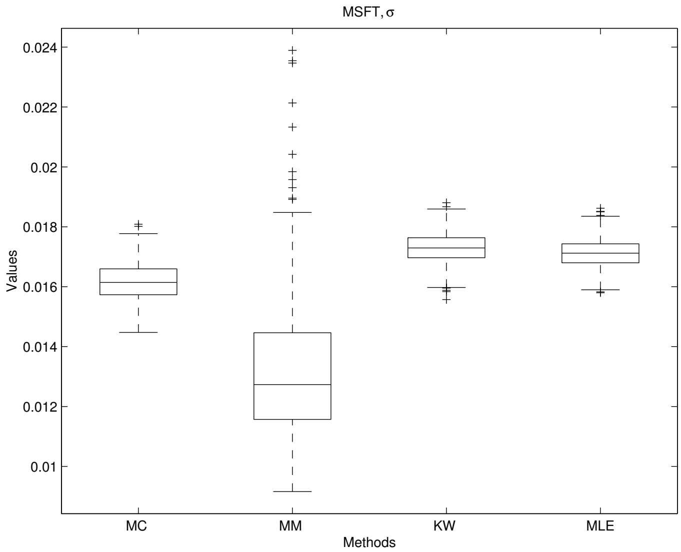


Figure 3: Boxplot diagrams of $\hat{\sigma}$, Microsoft

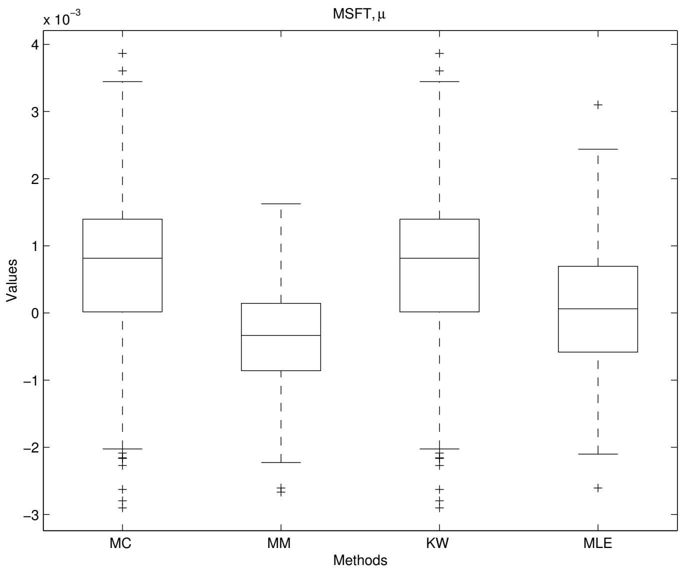


Figure 4: Boxplot diagrams of $\hat{\mu}$, Microsoft

	MC	MM	KW	MLE
mean of $\hat{\alpha}$	1.6458375	1.8295936	1.8491988	1.7967399
std of $\hat{\alpha}$	0.075413711	0.047666695	0.035276529	0.041418925
mean of $\hat{\beta}$	0.067143916	-1	0.47047319	0.12362177
std of $\hat{\beta}$	0.16323863	0	0.31157097	0.19035751
mean of $\hat{\sigma}$	0.016147967	0.013192684	0.017292292	0.01712232
std of $\hat{\sigma}$	0.000640467	0.002338541	0.00052722	0.000467339
mean of $\hat{\mu}$	0.000459517	-0.000357547	0.000709773	4.97935E-05
std of $\hat{\mu}$	0.00130509	0.00074554	0.001105434	0.000899346

Table V: Mean and standard deviation of estimated parameters according to the four methods, Microsoft

	MC	MM	KW	MLE	Normal Fit
KD	0.0466	0.1004	0.0588	0.0395	0.0509
ADS	0.0932	2.6111	0.1277	0.0791	27.7643

Table VI: Comparison in terms of goodness of fit measures defined in equations (20) and (21), Microsoft

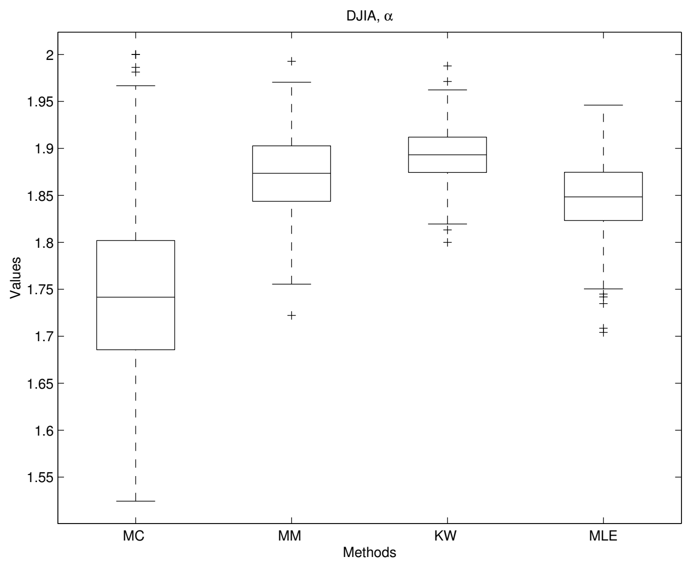


Figure 5: Boxplot diagrams of $\hat{\alpha}$, DJIA

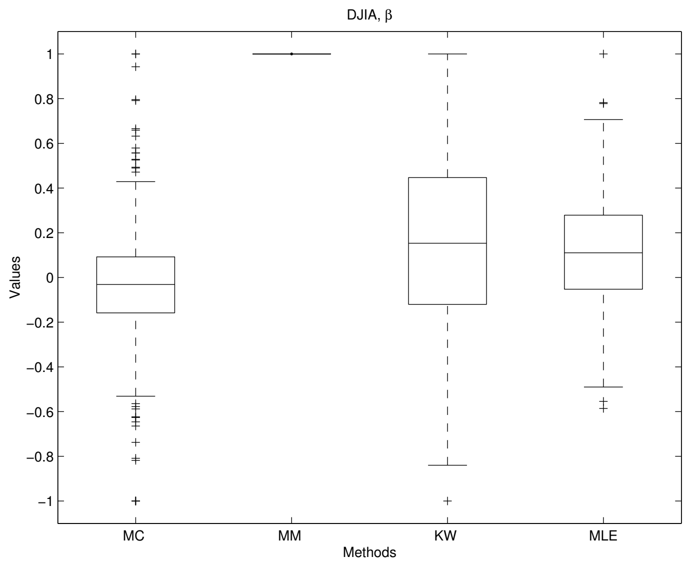


Figure 6: Boxplot diagrams of $\hat{\beta}$, DJIA

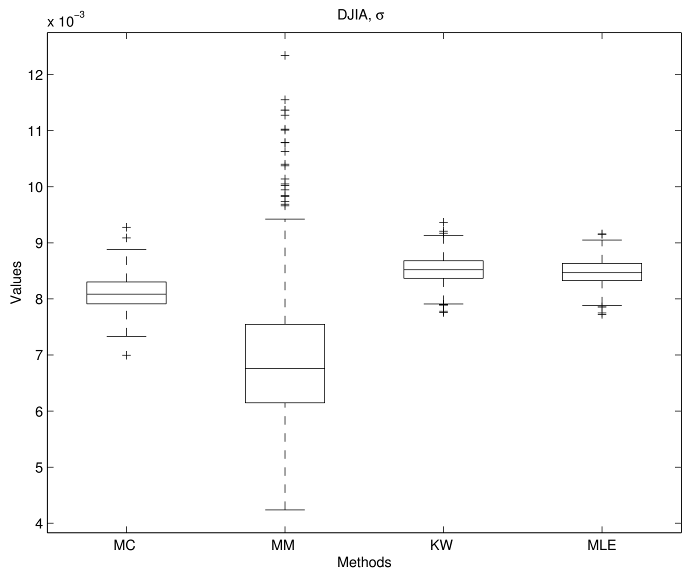


Figure 7: Boxplot diagrams of $\hat{\sigma}$, DJIA

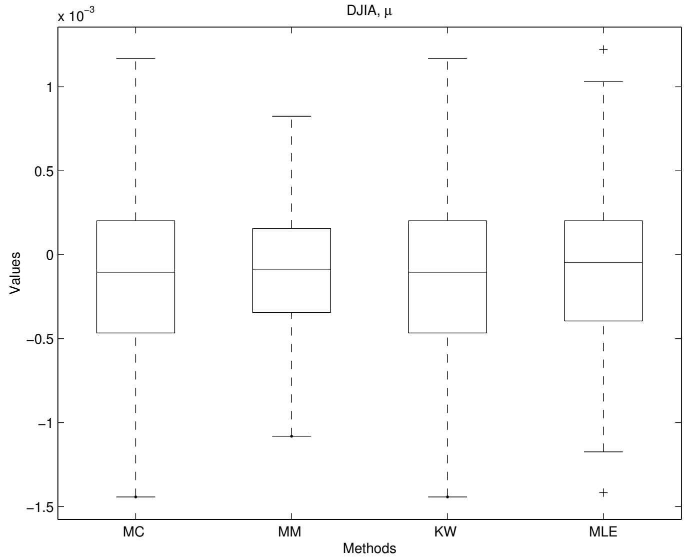


Figure 8: Boxplot diagrams of $\hat{\mu}$, DJIA

	MC	MM	KW	MLE
mean of $\hat{\alpha}$	1.7494998	1.8715545	1.8920675	1.847152
std of $\hat{\alpha}$	0.087647067	0.043739489	0.029119335	0.039937867
mean of $\hat{\beta}$	-0.034586045	1	0.15564377	0.11142357
std of $\hat{\beta}$	0.25759868	0	0.40483999	0.22890394
mean of $\hat{\sigma}$	0.008104254	0.006977522	0.008517372	0.008472094
std of $\hat{\sigma}$	0.000306711	0.00122087	0.000247869	0.00022901
mean of $\hat{\mu}$	-0.000382239	-9.37333E-05	-0.000104916	-9.09987E-05
std of $\hat{\mu}$	0.000546696	0.000352698	0.000484893	0.000439463

Table VII: Mean and standard deviation of estimated parameters according to the four methods, DJIA

	MC	MM	KW	MLE	Normal Fit
KD	0.0204	0.0790	0.0197	0.0180	0.0351
ADS	0.0612	2.0846	0.0563	0.0475	7.6447

Table VIII: Comparison in terms of goodness of fit measures defined in equations (20) and (21), DJIA

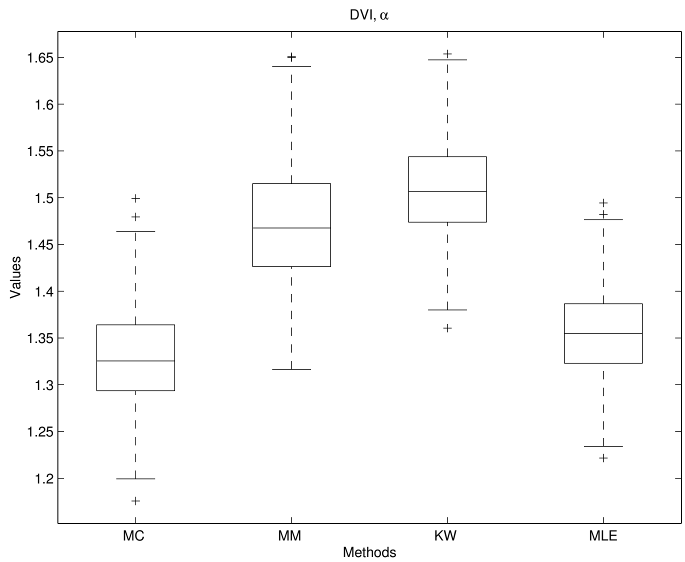


Figure 9: Boxplot diagrams of $\hat{\alpha}$, DVI Inc.

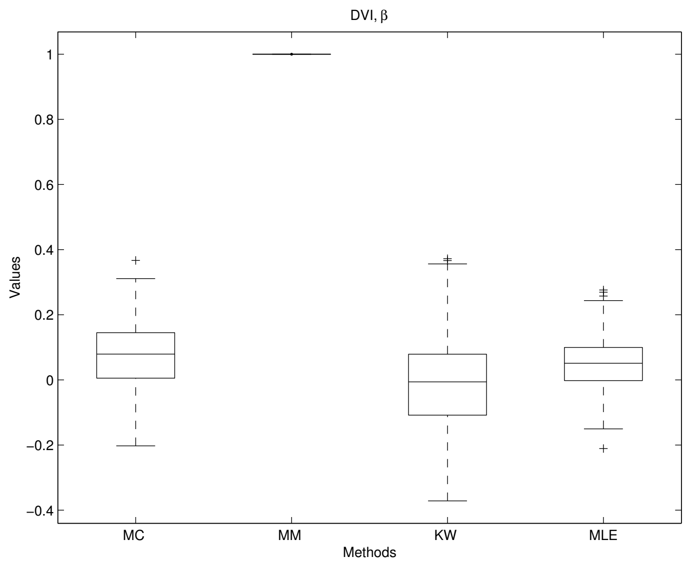


Figure 10: Boxplot diagrams of $\hat{\beta}$, DVI Inc.

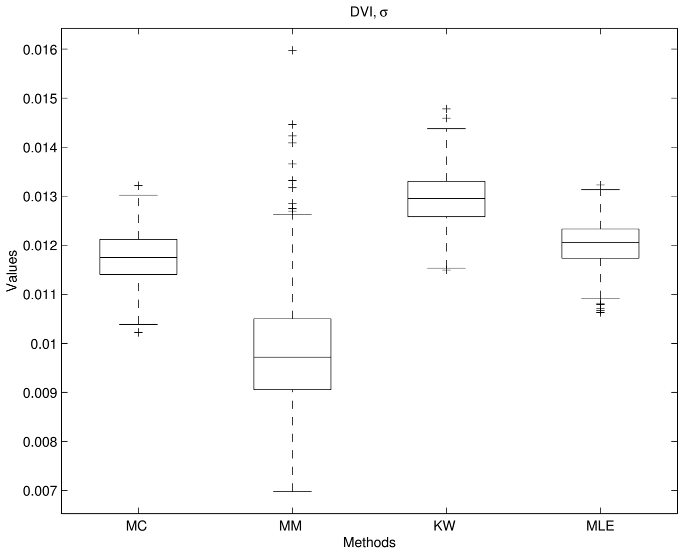


Figure 11: Boxplot diagrams of $\hat{\sigma}$, DVI Inc.

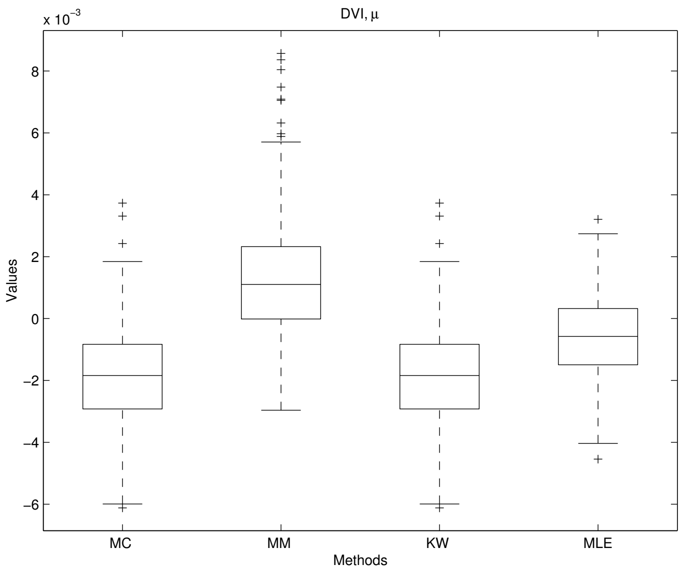


Figure 12: Boxplot diagrams of $\hat{\mu}$, DVI Inc.

	MC	MM	KW	MLE
mean of $\hat{\alpha}$	1.3283444	1.470059	1.5082385	1.355411
std of $\hat{\alpha}$	0.051544305	0.061008346	0.051359061	0.050175264
mean of $\hat{\beta}$	0.075472747	1	-0.008827654	0.048149033
std of $\hat{\beta}$	0.094929913	0	0.13790927	0.072749662
mean of $\hat{\sigma}$	0.011766864	0.009852498	0.012940463	0.012030631
std of $\hat{\sigma}$	0.000507936	0.001184099	0.00054783	0.000454509
mean of $\hat{\mu}$	0.001324993	0.001313229	-0.001833386	-0.000591671
std of $\hat{\mu}$	0.001972464	0.001859812	0.001567345	0.001289649

Table IX: Mean and standard deviation of estimated parameters according to the four methods, DVI Inc.

	MC	MM	KW	MLE	Normal Fit
KD	0.1177	0.1816	0.0670	0.0618	0.1261
ADS	0.2361	5.1388	0.1345	0.1238	6.2443e+008

Table X: Comparison in terms of goodness of fit measures defined in equations (20) and (21), DVI Inc.

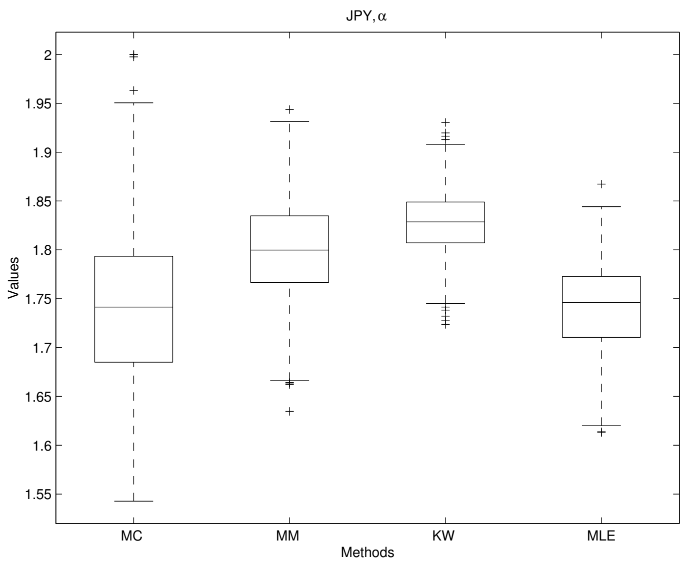


Figure 13: Boxplot diagrams of $\hat{\alpha}$, JPY/USD

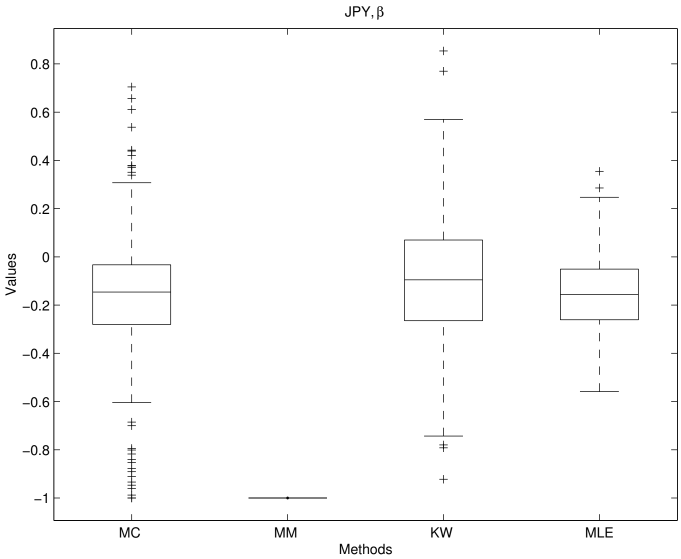


Figure 14: Boxplot diagrams of $\hat{\beta}$, JPY/USD

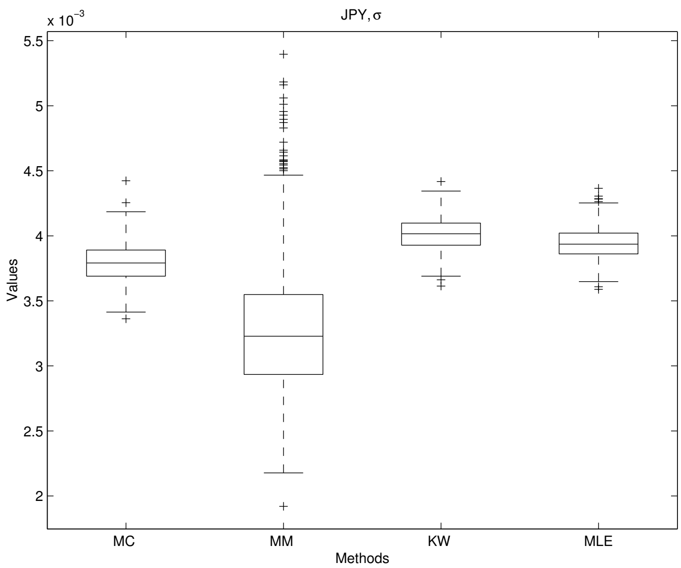


Figure 15: Boxplot diagrams of $\hat{\sigma}$, JPY/USD

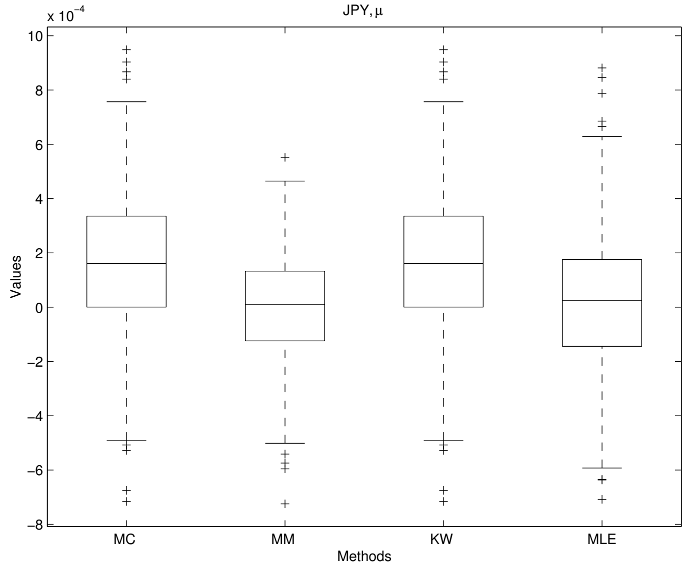


Figure 16: Boxplot diagrams of $\hat{\mu}$, JPY/USD

	MC	MM	KW	MLE
mean of $\hat{\alpha}$	1.7462556	1.8003459	1.8281921	1.7424175
std of $\hat{\alpha}$	0.086453937	0.050270346	0.032651054	0.044317015
mean of $\hat{\beta}$	-0.16562357	-1	-0.10156005	-0.15745804
std of $\hat{\beta}$	0.23872484	0	0.26965235	0.14908237
mean of $\hat{\sigma}$	0.003791973	0.003305307	0.004017066	0.00394396
std of $\hat{\sigma}$	0.000149153	0.000545953	0.000128849	0.000119034
mean of $\hat{\mu}$	-0.000146222	-1.72358E-06	0.00016135	1.61556E-05
std of $\hat{\mu}$	0.000239272	0.000206585	0.000256759	0.000248633

Table XI: Mean and standard deviation of estimated parameters according to the four methods, JPY/USD

	MC	MM	KW	MLE	Normal Fit
KD	0.1199	2.5404	0.1015	0.0831	0.0587
ADS	0.0461	0.1018	0.0302	0.0406	0.5171

Table XII: Comparison in terms of goodness of fit measures defined in equations (20) and (21), JPY/USD

4 Conclusion

On the basis of the empirical study, we can make the following observations regarding the estimators discussed. As far as the index of stability is concerned, the estimator of McCulloch (MC) has the worst performance in terms of the standard deviation of the bootstrapped estimates. For lower α, it produces better estimates for α, again, in terms of the standard deviation (see Figure 9). The last observation is in line with the Monte Carlo studies in [12]. The regression-type method of Kogon-Williams (KW) produces slightly less dispersed estimates of the index of stability than MLE in some cases, for instance Figures 1, 5 and 13. The latter outperforms all methods, and KW distinctly, in the estimation of β. All approaches, with the exception of the method of moments (MM), behave equally well in the estimation of the scale parameter. In contrast to its computational simplicity, MM has the worst overall performance in all cases (see Tables VI, VIII, X and XII) both in terms of the Kolmogorov distance and Anderson-Darling statistics. On the ground of the same criteria, MLE appears to be the most superior approach and the estimator of KW dominates that of MC in three cases (see Tables VIII, X and XII).

From computational viewpoint, the most complicated approach is MLE in which the speed of calculation strongly depends on the density approximation method. By far less computational burden is entailed by the estimator of Kogon-Williams followed by the quantile approach of McCulloch and the method of moments.

The goodness of fit measures imply that α-stable modeling prevails over the Gaussian modeling which is an expected result. The superiority of the former becomes particularly obvious when we compare the Anderson-Darling statistics, hence it follows that non-Gaussian α-stable laws significantly outperform the Gaussian distribution in forecasting extreme events, i.e. extreme losses or profits in financial context.

References

[1] R. J. Adler, R. Feldman and M. Taqqu, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhäuser, Boston, Basel, Berlin, 1998.

[2] V. Akgiray, C. G. Lamoureux, Estimation of Stable-law Parameter: A Comparative Study, *"Journal of Business and Economic Statistics"*, 7, 85-93, (1989).

[3] W. H. DuMouchel, On the Asymptotic Normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution, *"Annals of Statistics"*, 1, 948-957 (1973).

[4] E. Fama, The behavior of stock market prices, *"J. Bus. Univ. Chicago"*, 38, 34-105, (1965).

[5] E. Fama, R. Roll, Some Properties of Symmetric Stable Distributions, *"Journal of the American Statistical Association"*, 63, 817-836, (1968).

[6] E. Fama, R. Roll, Parameter Estimates for Symmetric Stable Distributions, *"Journal of the American Statistical Association"*, 66, 331-338, (1971) .

[7] I. A. Koutrouvelis, Regression-type Estimation of the Parameters of Stable Laws, *"Journal of the American Statistical Association"*, 75, 918-928, (1980).

[8] I. A. Koutrouvelis, An Iterative Procedure for the Estimation of the Parameters of Stable Laws, *"Communications in Statistics. Simulation and Computation,"* 10, 17-28, (1981).

[9] B. Mandelbrot, The variation of certain speculative prices, *"J. Bus. Univ. Chicago"*, 26, 394-419, (1963).

[10] J. H. McCulloch, Simple Consistent Estimators of Stable Distribution Parameters, *"Communications in Statistics. Simulation and Computation"*, 15, 1109-1136, (1986).

[11] S. J. Press, Estimation in Univariate and Multivariate Stable Distribution, *"Journal of the American Statistical Association"*, 67, 842-846, (1972b).

[12] S. T. Rachev and S. Mittnik *Stable Paretian Models in Finance*, John Wiley & Sons Ltd, 2000.

[13] G. Samorodnitsky, M. S. Taqqu, *Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance*, Chapman and Hall, New York, London, 1994.

[14] S. Stoyanov, B. Racheva-Iotova, Univariate stable laws in the field of finance - approximations of density and distribution functions, *"Journal of Concrete and Applicable Mathematics"*, to appear (2003).

[15] V. M. Zolotarev, *One-Dimensional Stable Distributions*, Nauka (in Russian), Moscow, 1983.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.4,397-417,2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

AN ANALOGUE OF HARDY'S THEOREM AND ITS L^p VERSION FOR THE DUNKL BESSEL TRANSFORM

Hatem MEJJAOLI and Khalifa TRIMÈCHE

Department of Mathematics

Faculty of Sciences of Tunis- Campus-

1060, Tunis-Tunisia

ABSTRACT

In this paper we study an analogue of Hardy's theorem and its L^p version for the Dunkl-Bessel transform on $\mathbb{R}_+^{d+1}$. More precisely for all $a > 0$, $b > 0$ and $p, q \in [1, +\infty]$, we determine the measurable functions f on $\mathbb{R}_+^{d+1}$ such that $e^{a||x||^2} f \in L^p_{k,\beta}(\mathbb{R}_+^{d+1})$ and $e^{b||y||^2}\mathcal{F}_{D,B}(f) \in L^p_{k,\beta}(\mathbb{R}_+^{d+1})$, where $L^p_{k,\beta}(\mathbb{R}_+^{d+1})$ designates the weighted Lebesgue space associated for the Dunkl-Bessel transform $\mathcal{F}_{D,B}$.

Key word; Dunkl-Bessel-transform; L^p version of Hardy's theorem
AMS subject classification: Primary 35R10; Secondary 44A15

1.INTRODUCTION

Hardy's theorem for the usual Fourier transform on $\mathbb{R}$ [8] asserts that f and its Fourier $\widehat{f}$ cannot both be very small. More precisely, let a and b be positive constants and assume that f is a measurable function on $\mathbb{R}$ such that $|f(x)| \leq Ce^{-ax^2}$ a.e and for all $y \in \mathbb{R}$, $|\hat{f}(y)| \leq C\, e^{-by^2}$ for some positive constant C. Then $f = 0$ a.e, if $ab > \frac{1}{4}$, $f = \text{const}\, e^{-ax^2}$ if $ab = \frac{1}{4}$, and there are infinitely many nonzero f if $ab < \frac{1}{4}$.

Considerable attention has been devoted recently to discovering generalizations to new contexts for Hardy's theorem. In particular this result have been obtained in [15] [6] for semisimple Lie groups and for the motion group. On the other hand M.G.Cowling and J.F.Price have studied in [1] an L^p version of Hardy's theorem which states that for $p, q \in [1, +\infty]$, at least one of them is finite, if $||e^{ax^2} f||_p < +\infty$ and $||e^{by^2}\hat{f}||_q < +\infty$ then $f = 0$ a.e, if $ab \geq \frac{1}{4}$. A generalization of this result to the Dunkl transform have been proved in [7].

In this paper we study an analogue of Hardy's and its L^p version for the Dunkl-Bessel transform $\mathcal{F}_{D,B}$ on $\mathbb{R}_+^{d+1}$.

The contents of the paper is as follows:

In the first section we recall the main results about the Dunkl operators, the Dunkl intertwining operator and its dual and the Dunkl transform.

We study in the second section the harmonic analysis associated to the Dunkl-Bessel operator. In particular we introduce and we study the Dunkl-Bessel intertwining operator $\mathcal{R}_{k,\beta}$ and we give properties of its dual ${}^t\mathcal{R}_{k,\beta}$ which plays an important role in the proofs of the main results of the paper.

The third section is devoted to Hardy's theorem and its L^p version. More precisely, we show that for all $p, q \in [1, +\infty]$, at least one of them is finite, if f is a measurable function on $\mathbb{R}_+^{d+1}$ such that $e^{a||x||^2} f \in L^p_{k,\beta}(\mathbb{R}_+^{d+1})$ and $e^{b||y||^2}\mathcal{F}_{D,B} f \in L^q_{k,\beta}(\mathbb{R}_+^{d+1})$ for some $a > 0$ and $b > 0$, then $f = 0$ a.e, if $ab \geq \frac{1}{4}$. For $p = q = +\infty$ we have $f = \mathrm{const}\, e^{-a||x||^2}$ if $ab = \frac{1}{4}$, and for all p, q there are infinitely many nonzero f if $ab < \frac{1}{4}$.

2.HARMONIC ANALYSIS ASSOCIATED WITH THE DUNKL OPERATOR

In this section we collect some notations and results on Dunkl operators, the Dunkl intertwining operator and its dual and the Dunkl transform (see [3], [4], [5]).

2.1. Reflection groups, root system and multiplicity functions

We consider $\mathbb{R}^d$ with the euclidean scalar product $\langle , \rangle$ and $||x|| = \sqrt{\langle x, x\rangle}$. On $\mathbb{C}^d$, $||.||$ denotes also the standard Hermitian norm, while $\langle z, w\rangle = \sum_{j=1}^{d} z_j \overline{w_j}$.

For $\alpha \in \mathbb{R}^d \backslash \{0\}$, let σ_α be the reflection in the hyperplan $H_\alpha \subset \mathbb{R}^d$ orthogonal to α, i.e.

$$\sigma_\alpha(x) = x - 2\frac{\langle \alpha, x\rangle}{||\alpha||^2}\alpha. \tag{1}$$

A finite set $R \subset \mathbb{R}^d \backslash \{0\}$ is called a root system if $R \cap \mathbb{R}^d.\alpha = \{\alpha, -\alpha\}$ and $\sigma_\alpha R = R$ for all $\alpha \in R$. For a given root system R the reflection $\sigma_\alpha, \alpha \in R$, generate a finite group $W \subset O(d)$, called the reflection group associated with R . All reflections in W correspond to suitable pairs of roots. For a given $\beta \in \mathbb{R}^d \backslash \cup_{\alpha\in R} H_\alpha$, we fix the positive subsystem $R_+ = \{\alpha \in R \,/ \langle \alpha, \beta\rangle > 0\}$, then for each $\alpha \in R$ either $\alpha \in R_+$ or $-\alpha \in R_+$.

A function $k : R \longrightarrow \mathbb{C}$ on a root system R is called a multiplicity function if it is invariant under the action of the associated reflection group W. If one regards k as a function on the corresponding reflections, this means that k is constant on the conjugacy classes of reflections in W.

For abbreviation, we introduce the index

$$\gamma = \gamma(k) = \sum_{\alpha\in R_+} k(\alpha). \tag{2}$$

Moreover, let ω_k denotes the weight function

$$\omega_k(x) = \prod_{\alpha\in R_+} |\langle \alpha, x\rangle|^{2k(\alpha)}, \tag{3}$$

which is invariant and homogeneous of degree 2γ.

For $d = 1$ and $W = \mathbf{Z}_2$, the multiplicity function k is a single parameter denoted $\gamma > 0$, and

$$\forall\, x \in I\!R,\ \omega_k(x) = |x|^{2\gamma}. \tag{4}$$

We introduce the Mehta-type constant

$$c_k = (\int_{I\!R^d} \exp(-||x||^2)\omega_k(x)\, dx)^{-1}. \tag{5}$$

2.2.Dunkl operators and Dunkl intertwining operator and its dual

Notations. We denote by
- $C(I\!R^d)(resp\ C_c(I\!R^d))$ the space of continuous functions on $I\!R^d$ (resp. with compact support).
- $\mathcal{E}(I\!R^d)$ the space of C^∞-functions on $I\!R^d$.
- $\mathcal{S}(I\!R^d)$ the space of C^∞-functions on $I\!R^d$ which are rapidly decreasing as their derivatives.
- $D(I\!R^d)$ the space of C^∞-functions on $I\!R^d$ which are of compact support.

We provide these spaces with the classical topology .

The Dunkl operators $T_j,\ j\ = 1\ ,...,\ d$, on $I\!R^d$ associated with the finite reflection group W and multiplicity function k are given by

$$T_j f(x) = \frac{\partial}{\partial x_j} f(x) + \sum_{\alpha\in R_+} k(\alpha)\alpha_j \frac{f(x) - f(\sigma_\alpha(x))}{\langle \alpha, x\rangle}, \quad f\ \in\ C^1(I\!R^d). \tag{6}$$

For all f, g in $\mathcal{E}(I\!R^d)$ with g is W-invariant, we have the product rule

$$T_j(fg) = (T_j f)g + f(T_j g), \quad j = 1, ..., d. \tag{7}$$

In the case $k = 0$, the T_j, $j = 1, ..., d$, reduce to the corresponding partial derivatives. In this paper, we will assume throughout that $k \geq 0$ and $\gamma > 0$.

We define the Dunkl-Laplace operator on $I\!R^d$ by

$$\triangle_k f(x) = \sum_{j=1}^{d} T_j^2 f(x) = \triangle_d f(x) + 2\sum_{\alpha\in R^+} k(\alpha)[\frac{\langle \nabla f(x), \alpha\rangle}{\langle \alpha, x\rangle} - \frac{f(x) - f(\sigma_\alpha(x))}{\langle \alpha, x\rangle^2}], \tag{8}$$

where $\triangle_d = \sum_{j=1}^{d} \partial_j^2$ and ∇ are respectively the Laplacian and the gradient on $I\!R^d$.

For $y \in I\!R^d$, the system

$$\begin{cases} T_j u(x,y) &= y_j u(x,y), \quad j = 1,...,d, \\ u(0,y) &= 1, \qquad\qquad \text{for all y} \in I\!R^d. \end{cases}$$

admits a unique analytic solution on $I\!R^d$, which will be denoted by $K(x,y)$ and called Dunkl kernel. This kernel has a unique holomorphic extension to $\mathbb{C}^d \times \mathbb{C}^d$.

Examples.

1) If $d = 1$ and $W = Z_2$, the Dunkl kernel is given by

$$K(z,t) = j_{\gamma-\frac{1}{2}}(izt) + \frac{zt}{2\gamma+1} j_{\gamma+\frac{1}{2}}(izt), \quad z,\ t \in \mathbb{C}, \tag{9}$$

where for $\alpha \geq \frac{-1}{2}$, j_α is the normalized Bessel function defined by

$$j_\alpha(z) = 2^\alpha \Gamma(\alpha+1) \frac{J_\alpha(z)}{z^\alpha} = \Gamma(\alpha+1) \sum_{n=0}^{\infty} \frac{(-1)^n (\frac{z}{2})^{2n}}{n! \Gamma(\alpha+n+1)}, \tag{10}$$

with J_α is the Bessel function of first kind and index α.

2) The Dunkl kernel of index $\gamma = \sum_{l=1}^{d} \alpha_l,\ \alpha_l > 0$, associated with the reflection group $\mathbf{Z}_2 \times ... \times \mathbf{Z}_2$ on $I\!R^d$ is given for all $x, y \in I\!R^d$ by

$$K(x,y) = \prod_{l=1}^{d} K(x_l, y_l), \tag{11}$$

where $K(x_l, y_l)$ is the function defined by Eq.(9).

The Dunkl kernel possesses the following properties:

i) For every $z, t \in \mathbb{C}^d$, we have

$$K(z,t) = K(t,z) \quad ; K(z,0) = 1 \quad \text{and} \quad K(\lambda z, t) = K(z, \lambda t), \ \text{for all } \lambda \in \mathbb{C}. \tag{12}$$

ii) For all $\nu \in I\!N^d, x \in I\!R^d$ and $z \in \mathbb{C}^d$, we have

$$|D_z^\nu K(x,z)| \leq ||x||^{|\nu|} \exp(||x|| ||Rez||), \tag{13}$$

and for all $x, y \in I\!R^d$:

$$|K(ix,y)| \leq 1, \tag{14}$$

with $D_z^\nu = \frac{\partial^\nu}{\partial z_1^{\nu_1} ... \partial z_d^{\nu_d}}$ and $|\nu| = \nu_1 + ... + \nu_d$.

iii) For all $x, y \in I\!R^d$ and $w \in W$ we have

$$K(-ix,y) = \overline{K(ix,y)}, \quad \text{and} \quad K(wx, wy) = K(x,y). \tag{15}$$

$i\nu$) The function $K(x,z)$ admits for all $x \in I\!R^d$ and $z \in \mathbb{C}^d$ the following Laplace type integral representation

$$K(x,z) = \int_{I\!R^d} e^{<y,z>} d\mu_x(y), \tag{16}$$

where $d\mu_x$ is a probability measure on $I\!R^d$, with support in the closed ball $B(o, ||x||)$ of center o and radius $||x||$.

The Dunkl intertwining operator V_k is defined on $C(I\!R^d)$ by

$$\forall x \in I\!R^d, \quad V_k f(x) = \int_{I\!R^d} f(y) d\mu_x(y), \tag{17}$$

where $d\mu_x$ is the measure given by the relation (16).

We have

$$\forall x \in I\!R^d, \;\; \forall z \in \mathbb{C}^d, \;\; K(x,z) = V_k(e^{\langle \cdot, z\rangle})(x).$$

The operator ${}^tV_k$ satisfying for f in $D(I\!R^d)$ and g in $C(I\!R^d)$ the relation

$$\int_{I\!R^d} {}^tV_k(f)(y) g(y) dy = \int_{I\!R^d} V_k(f)(x) g(x) \omega_k(x) dx, \tag{18}$$

is given by

$${}^tV_k(f)(y) = \int_{I\!R^d} f(x) d\nu_y(x), \tag{19}$$

where ν_y is a positive measure on $I\!R^d$ with support in the set $\{x \in I\!R^d, ||x|| \geq ||y||\}$. This operator is called the dual Dunkl intertwining operator.

The operators V_k and ${}^tV_k$ satisfy the following properties:

i)The operator V_k is a topological isomorphism from $\mathcal{E}(I\!R^d)$ onto itself satisfying the transmutation relation

$$\forall x \in I\!R^d, \quad T_j V_k(f)(x) = V_k(\frac{\partial}{\partial y_j} f)(x), \quad j = 1, ..., d, f \in \mathcal{E}(I\!R^d). \tag{20}$$

ii) The operator ${}^tV_k$ is a topological isomorphism from $D(I\!R^d)$(resp. $\mathcal{S}(I\!R^d)$) onto itself, satisfying the transmutation relations

$$\forall y \in I\!R^d, \; {}^tV_k(T_j f)(y) = \frac{\partial}{\partial y_j} {}^tV_k(f)(y), j = 1, ..., d, \; f \in D(I\!R^d). \tag{21}$$

We denote by $L^p_k(I\!R^d)$ the space of measurable functions on $I\!R^d$ such that

$$||f||_{k,p} \;\; = \;\; (\int_{I\!R^d} |f(x)|^p \omega_k(x) \; dx)^{\frac{1}{p}} < +\infty, \quad \text{if } 1 \leq p < +\infty.$$

Theorem 2.2.1. *Let $(\nu_y)_{y \in I\!R^d}$ be the family of measures defined by Eq.(19) and let $f \in L^1_k(I\!R^d)$. Then for almost all y (with respect to Lebesgue measure on $I\!R^d$), f is ν_y-integrable, the function*

$$y \mapsto \int_{I\!R^d} f(x) d\nu_y(x),$$

which will also be denoted by ${}^tV_k(f)$*, is defined almost everywhere on* $I\!R^d$ *and is Lebesgue integrable. Moreover for all bounded continuous function* g *on* $I\!R^d$*, we have the formula*

$$\int_{I\!R^d} {}^tV_k(f)(y)g(y)dy = \int_{I\!R^d} V_k(f)(x)g(x)\omega_k(x)dx. \tag{22}$$

Examples:

1) When $d = 1$ and $W = \mathbf{Z}_2$, for all $x \in I\!R\backslash\{0\}$ the measure μ_x of the Dunkl intertwining operator V_k defined by Eq.(17) is given by $d\mu_x(y) = \mathcal{K}(x,y)dy$, with

$$\mathcal{K}(x,y) = \frac{\Gamma(\gamma+\frac{1}{2})}{\sqrt{\pi}\Gamma(\gamma)}|x|^{-2\gamma}(|x|-y)^{\gamma-1}(|x|+y)^{\gamma}1_{]-|x|,|x|[}(y), \tag{23}$$

where $1_{]-|x|,|x|[}$ is the characteristic function of the interval $]-|x|,|x|[$.

The dual Dunkl intertwining operator ${}^tV_k$ is defined by Eq.(18) with for all $y \in I\!R$ we have $d\nu_y(x) = \mathcal{K}(x,y)\omega_k(x)dx$, where $\mathcal{K}$ is given by Eq.(23).

2) The Dunkl intertwining operator V_k of index $\gamma = \sum_{l=1}^{d}\alpha_l$, $\alpha_l > 0$, associated with the reflection group $\mathbf{Z}_2 \times ... \times \mathbf{Z}_2$ on $I\!R^d$, is given for all f in $\mathcal{E}(I\!R^d)$ and $x \in I\!R^d\backslash \bigcup_{l=1}^{d} H_l$, with $H_l = \{x \in I\!R^d / x_l = 0\}$ by

$$V_k(f)(x) = \int_{I\!R^d}\mathcal{K}(x,y)f(y)dy,$$

where

$$\mathcal{K}(x,y) = \prod_{l=1}^{d}\mathcal{K}(x_l,y_l),$$

with $\mathcal{K}(x_l,y_l)$ is given by the relation (23). By change of variables we obtain

$$\begin{aligned}\forall\, x \in I\!R^d,\ V_k(f)(x) \;&=\; [\prod_{l=1}^{d}\frac{\Gamma(\alpha_l+\frac{1}{2})}{\sqrt{\pi}\Gamma(\alpha_l)}]\int_{[-1,1]^d} f(t_1x_1,t_2x_2,...,t_dx_d)\\ &\quad\times\left[\prod_{l=1}^{d}(1-t_l)^{\alpha_l-1}(1+t_l)^{\alpha_l}\right]dt_1...dt_d.\end{aligned} \tag{24}$$

It can also be written in the form

$$\forall\, x \in I\!R^d,\ V_kf(x) = (V_k)^1 \times ... \times (V_k)^d(f)(x), \tag{25}$$

where for all g in $\mathcal{E}(I\!R)$ we have

$$\forall\, x \in I\!R^d,\ (V_k)^l g(x) = \frac{2\Gamma(\alpha_l+\frac{1}{2})}{\sqrt{\pi}\Gamma(\alpha_l)}\int_{-1}^{1} g(t_lx_l)(1-t_l)^{\alpha_l-1}(1+t_l)^{\alpha_l}dt_l.$$

(See [17]).

2.3. The Dunkl transform

The Dunkl transform of a function f in $D(I\!R^d)$ is given by

$$\forall y \in I\!R^d, \quad \mathcal{F}_D(f)(y) = \int_{I\!R^d} f(x)K(-iy, x)\omega_k(x)dx. \tag{26}$$

It satisfies the following properties:

i) For f in $L^1_k(I\!R^d)$ we have

$$||\mathcal{F}_D(f)||_{k,\infty} \leq ||f||_{k,1}. \tag{27}$$

ii) For all f in $\mathcal{S}(I\!R^d)$ we have

$$\mathcal{F}_D(f) = \mathcal{F} \circ {}^tV_k(f), \tag{28}$$

where $\mathcal{F}$ is the classical Fourier transform on $I\!R^d$.

iii) For f in $\mathcal{S}(I\!R^d)$ we have

$$\forall y \in I\!R^d, \quad \mathcal{F}_D(T_j f)(y) = iy_j \mathcal{F}_D(f)y), \quad j = 1, ..., d. \tag{29}$$

$i\nu$) For all f in $L^1_k(I\!R^d)$ such that $\mathcal{F}_D(f)$ is in $L^1_k(I\!R^d)$, we have the inversion formula

$$f(y) = \frac{c_k^2}{4^{\gamma+\frac{d}{2}}} \int_{I\!R^d} \mathcal{F}_D(f)(x)K(ix, y)\omega_k(x)\, dx, \quad a.e. \tag{30}$$

Theorem 2.3.1. *The Dunkl transform $\mathcal{F}_D$ is a topological isomorphism:*

i) From $\mathcal{S}(I\!R^d)$ onto itself.

ii) From $D(I\!R^d)$ onto $H(\mathbb{C}^d)$ (the space of entire functions on $\mathbb{C}^d$, rapidly decreasing of exponential type.)

The inverse transform $\mathcal{F}_D^{-1}$ is given by

$$\forall y \in I\!R^d, \quad \mathcal{F}_D^{-1}(f)(y) = \frac{c_k^2}{4^{\gamma+\frac{d}{2}}} \mathcal{F}_D(f)(-y), \quad f \in S(I\!R^d). \tag{31}$$

Theorem 2.3.2.i)*(Plancherel formula for $\mathcal{F}_D$) For all f in $\mathcal{S}(I\!R^d)$ we have*

$$\int_{I\!R^d} |f(x)|^2 \omega_k(x)\, dx = \frac{c_k^2}{4^{\gamma+\frac{d}{2}}} \int_{I\!R^d} |\mathcal{F}_D(f)(\xi)|^2 \omega_k(\xi)\, d\xi. \tag{32}$$

ii) (Plancherel theorem for $\mathcal{F}_D$) The renormalized Dunkl transform $f \to 2^{-(\gamma+\frac{d}{2})} c_k \mathcal{F}_D(f)$ can be uniquely extended to an isometric isomorphism on $L^2_k(I\!R^d)$.

3. HARMONIC ANALYSIS ASSOCIATED WITH THE DUNKL-BESSEL-LAPLACE OPERATOR

Notations. We denote by
-$\mathbb{R}_+^{d+1} = \mathbb{R}^d \times [0, +\infty[$.
- $x = (x_1, ..., x_d, x_{d+1}) = (x', x_{d+1}) \in \mathbb{R}_+^{d+1}$.
- $C_*(\mathbb{R}^{d+1})(resp.\ C_{*,c}(\mathbb{R}^{d+1}))$ the space of continuous functions on $\mathbb{R}^{d+1}$ (resp. with compact support), even with respect to the last variable.
- $C_*^p(\mathbb{R}^{d+1})(resp.\ C_{*,c}^p(\mathbb{R}^{d+1}))$ the space of functions of class C^p on $\mathbb{R}^{d+1}$, even with respect to the last variable .
-$\mathcal{E}_*(\mathbb{R}^{d+1})$ (resp. $D_*(\mathbb{R}^{d+1})$) the space of C^∞-functions on $\mathbb{R}^{d+1}$ (resp. with compact support), even with respect to the last variable.
- $\mathcal{S}_*(\mathbb{R}^{d+1})$ the Schwartz space consisting of functions on $\mathbb{R}^{d+1}$ which are even with respect to the last variable.

We provide these spaces with the classical topology.

3.1. The Dunkl-Bessel-Laplace operator and the Dunkl-Bessel intertwining operator

We consider the Dunkl-Bessel-Laplace operator $\triangle_{k,\beta}$ defined by $\forall\, x = (x', x_{d+1}) \in \mathbb{R}^d \times]0, +\infty[$,

$$\triangle_{k,\beta} f(x) = \triangle_{k,x'} f(x', x_{d+1}) + L_{\beta,x_{d+1}} f(x', x_{d+1}),\ f \in C_*^2(\mathbb{R}^{d+1}), \tag{33}$$

where $\triangle_k$ is the Dunkl-Laplace operator on $\mathbb{R}^d$, and L_β the Bessel operator on $]0, +\infty[$ given by

$$L_\beta = \frac{d^2}{dx_{d+1}^2} + \frac{2\beta+1}{x_{d+1}} \frac{d}{dx_{d+1}}, \quad \beta > -\frac{1}{2}.$$

The function Λ given by

$$\Lambda(x, z) = K(-ix', z') j_\beta(x_{d+1} z_{d+1}),\ (x, z) \in \mathbb{R}_+^{d+1} \times \mathbb{C}^{d+1}, \tag{34}$$

satisfies the following properties:

i) For every $z, t \in \mathbb{C}^{d+1}$, we have

$$\Lambda(z, t) = \Lambda(t, z) \quad ; \Lambda(z, 0) = 1 \quad \text{and} \quad \Lambda(\lambda z, t) = \Lambda(z, \lambda t), \quad \text{for all } \lambda \in \mathbb{C}. \tag{35}$$

ii) For all $\nu \in \mathbb{N}^d, x \in \mathbb{R}_+^{d+1}$ and $z \in \mathbb{C}^{d+1}$, we have

$$|D_z^\nu \Lambda(x, z)| \leq ||x||^{|\nu|} \exp(||x|| ||Im z||), \tag{36}$$

and for all $x, y \in \mathbb{R}_+^{d+1}$:

$$|\Lambda(x, y)| \leq 1, \tag{37}$$

with $D_z^\nu = \frac{\partial^\nu}{\partial z_1^{\nu_1}...\partial z_{d+1}^{\nu_{d+1}}}$ and $|\nu| = \nu_1 + ... + \nu_{d+1}$.

On the other hand the function Λ is a solution of the system

$$\begin{cases} (\triangle_{k,\beta})_x u(x,z) = -||z||^2 u(x,z), & (x,z) \in \mathbb{R}_+^{d+1} \times \mathbb{C}^{d+1} \\ u(0,z) = 1; & \frac{\partial}{\partial x_{d+1}} u((x',0),z) = 0, \text{for all } z \in \mathbb{C}^{d+1}. \end{cases} \tag{38}$$

The Dunkl-Bessel intertwining operator is the operator defined on $C_*(\mathbb{R}^{d+1})$ by

$$\mathcal{R}_{k,\beta} f(x', x_{d+1}) = \begin{cases} \frac{2\Gamma(\beta+1)}{\sqrt{\pi}\Gamma(\beta+\frac{1}{2})} x_{d+1}^{-2\beta} \int_0^{x_{d+1}} (x_{d+1}^2 - t^2)^{\beta-\frac{1}{2}} V_k f(x', t) dt, & x_{d+1} > 0, \\ V_k f(x', 0), & x_{d+1} = 0, \end{cases} \tag{39}$$

where V_k is the Dunkl intertwining operator given by Eq.(17).

The operator $\mathcal{R}_{k,\beta}$ can also be write in the form

$$\mathcal{R}_{k,\beta} = V_k \otimes \mathcal{R}_\beta, \tag{40}$$

where $\mathcal{R}_\beta$ is the Riemann-Liouville integral operator given for all even continuous functions g on $\mathbb{R}$ by

$$\forall t > 0,\ \mathcal{R}_\beta(g)(t) = \frac{2\Gamma(\beta+1)}{\sqrt{\pi}\Gamma(\beta+\frac{1}{2})} t^{-2\beta} \int_0^t (t^2 - y^2)^{\beta-\frac{1}{2}} g(y) dy.$$

It satisfies the transmutation relation

$$\mathcal{R}_\beta \circ \frac{d^2}{dy^2} = L_\beta \circ \mathcal{R}_\beta. \tag{41}$$

(See [16]).

We remark that

$$\forall\, (x,z) \in \mathbb{R}_+^{d+1} \times \mathbb{C}^{d+1}, \quad \Lambda(x,z) = \mathcal{R}_{k,\beta}(e^{-i\langle .,z'\rangle} \cos(z_{d+1}.))(x). \tag{42}$$

We define for all $x \in \mathbb{R}_+^{d+1}$ the measure $\zeta_x^{k,\beta}$ by

$$d\zeta_x^{k,\beta}(y) = \frac{2\Gamma(\beta+1)}{\sqrt{\pi}\Gamma(\beta+\frac{1}{2})} x_{d+1}^{-2\beta} (x_{d+1}^2 - y_{d+1}^2)^{\beta-\frac{1}{2}} 1_{]0,x_{d+1}[}(y_{d+1}) d\mu_{x'}(y') dy_{d+1}, \tag{43}$$

where $d\mu_{x'}$ is the measure given by Eq.(17) and $1_{]0,x_{d+1}[}$ is the characteristic function of the interval $]0, x_{d+1}[$.

Hence from Eq.(39) the operator $\mathcal{R}_{k,\beta}$ can also be written in the form

$$\mathcal{R}_{k,\beta}f(x) = \int_{I\!R_+^{d+1}} f(y)d\zeta_x^{k,\beta}(y). \tag{44}$$

It is a topological isomorphism from $\mathcal{E}_*(I\!R^{d+1})$ onto itself satisfying

$$\forall\, f \in \mathcal{E}_*(I\!R^{d+1}),\ \triangle_{k,\beta}\mathcal{R}_{k,\beta}f = \mathcal{R}_{k,\beta}\triangle_{d+1}f, \tag{45}$$

where $\triangle_{d+1} = \sum_{j=1}^{d+1}\partial_j^2$.

Example:

We consider the reflection group $\mathbf{Z}_2 \times ... \times \mathbf{Z}_2$. The Dunkl-Bessel intertwining operator $\mathcal{R}_{k,\beta}$ on $I\!R_+^{d+1}$, is given for all f in $\mathcal{E}(I\!R_*^{d+1})$ and $x \in I\!R_+^{d+1}$ by

$$\begin{aligned}\mathcal{R}_{k,\beta}(f)(x) &= [\prod_{l=1}^{d}\frac{\Gamma(\alpha_l+\frac{1}{2})}{\pi^{\frac{d+1}{2}}\Gamma(\alpha_l)}\frac{\Gamma(\beta+1)}{\Gamma(\beta+\frac{1}{2})}]\int_{[-1,1]^d}\int_0^1 f(t_1x_1, t_2x_2, ..., t_dx_d, t_{d+1}x_{d+1})\\ &\quad\times[\prod_{l=1}^{d}(1-t_l)^{\alpha_l-1}(1+t_l)^{\alpha_l}](1-t_{d+1}^2)^{\beta-\frac{1}{2}}dt_1...dt_ddt_{d+1}.\end{aligned}$$

3.2. The dual of the Dunkl-Bessel intertwining operator

The dual of the Dunkl-Bessel intertwining operator $\mathcal{R}_{k,\beta}$ is the operator ${}^t\mathcal{R}_{k,\beta}$ defined on $D_*(I\!R^{d+1})$ by : $\forall y = (y', y_{d+1}) \in I\!R^d \times [0, \infty[$,

$${}^t\mathcal{R}_{k,\beta}(f)(y', y_{d+1}) = \frac{2\Gamma(\beta+1)}{\sqrt{\pi}\Gamma(\beta+\frac{1}{2})}\int_{y_{d+1}}^{\infty}(s^2 - y_{d+1}^2)^{\beta-\frac{1}{2}}\ {}^tV_kf(y', s)sds, \tag{46}$$

where ${}^tV_k$ is the dual Dunkl intertwining operator given by Eq.(18).
We can write ${}^t\mathcal{R}_{k,\beta}$ in the form

$${}^t\mathcal{R}_{k,\beta} = {}^tV_k \otimes {}^t\mathcal{R}_\beta, \tag{47}$$

where ${}^t\mathcal{R}_\beta$ is the Weyl integral operator defined for all even continuous function g on $I\!R$ with compact support, by

$$\forall\, y \geq 0,\ {}^t\mathcal{R}_\beta(f)(y) = \frac{2\Gamma(\beta+1)}{\sqrt{\pi}\Gamma(\beta+\frac{1}{2})}\int_y^{\infty}(s^2 - y^2)^{\beta-\frac{1}{2}}f(s)sds.$$

(See [16]).

We define for all $y \in I\!R_+^{d+1}$ the measure $\varrho_y^{k,\beta}$ by

$$d\varrho_y^{k,\beta}(x) = \frac{2\Gamma(\beta+1)}{\sqrt{\pi}\Gamma(\beta+\frac{1}{2})}(x_{d+1}^2 - y_{d+1}^2)^{\beta-\frac{1}{2}}x_{d+1}1_{]y_{d+1},+\infty[}(x_{d+1})d\nu_{y'}(x')dx_{d+1}, \tag{48}$$

where $d\nu_{y'}$ is the measure given by Eq.(19) and $1_{]y_{d+1},+\infty[}$ is the characteristic function of the interval $]y_{d+1},+\infty[$.

Hence from Eq.(46) the operator ${}^t\mathcal{R}_{k,\beta}$ can also be written in the form

$$ {}^t\mathcal{R}_{k,\beta}(f)(y) = \int_{\mathbb{R}_+^{d+1}} f(x) d\varrho_y^{k,\beta}(x). \qquad (49)$$

It satisfies for f in $D_*(\mathbb{R}^{d+1})$ and g in $\mathcal{E}_*(\mathbb{R}^{d+1})$ the following relation

$$\int_{\mathbb{R}_+^{d+1}} {}^t\mathcal{R}_{k,\beta}(f)(y)g(y)d\mu_{k,\beta}(y) = \int_{\mathbb{R}_+^{d+1}} f(y)\mathcal{R}_{k,\beta}(g)(y)dy. \qquad (50)$$

Moreover it is a topological isomorphism from $D_*(\mathbb{R}^{d+1})$ (resp. $\mathcal{S}_*(\mathbb{R}^{d+1})$) onto itself.

Theorem 3.2.1. *Let $(\varrho_y^{k,\beta})_{y\in\mathbb{R}_+^{d+1}}$ be the family of measures defined by Eq.(49) and let $f \in L^1_{k,\beta}(\mathbb{R}_+^{d+1})$. Then for almost all y (with respect to the Lebesgue measure on $\mathbb{R}_+^{d+1}$), f is $\varrho_y^{k,\beta}$-integrable, the function*

$$y \mapsto \int_{\mathbb{R}_+^{d+1}} f(x) d\varrho_y^{k,\beta}(x),$$

which will also be denoted by ${}^t\mathcal{R}_{k,\beta}(f)$, is defined almost everywhere on $\mathbb{R}_+^{d+1}$, and is Lebesgue integrable. Moreover for all bounded continuous function g on $\mathbb{R}^{d+1}$, we have the formula

$$\int_{\mathbb{R}_+^{d+1}} {}^t\mathcal{R}_{k,\beta}(f)(y)g(y)dy = \int_{\mathbb{R}_+^{d+1}} f(x)\, \mathcal{R}_{k,\beta}g(x)d\mu_{k,\beta}(x). \qquad (51)$$

Proof

We deduce the result by using the relations (47), (48) and Theorem 2.2.1.

3.3. The Dunkl-Bessel transform

Notations. We denote by $L^p_{k,\beta}(\mathbb{R}_+^{d+1})$ the space of measurable functions on $\mathbb{R}_+^{d+1}$ such that

$$||f||_{k,\beta,p} = \left(\int_{\mathbb{R}_+^{d+1}} |f(x)|^p d\mu_{k,\beta}(x)\, dx\right)^{\frac{1}{p}} < +\infty, \quad \text{if } 1 \le p < +\infty,$$

$$||f||_{k,\beta,\infty} = ess\ sup_{x\in\mathbb{R}_+^{d+1}}|f(x)| < +\infty,$$

where $d\mu_{k,\beta}$ is the measure given by

$$d\mu_{k,\beta}(x', x_{d+1}) = \omega_k(x')x_{d+1}^{2\beta+1}dx'dx_{d+1}.$$

Definition 3.3.1. The Dunkl-Bessel transform is given for f in $D_*(\mathbb{R}^{d+1})$ by
$\forall\, y = (y', y_{d+1}) \in \mathbb{R}_+^{d+1}$,

$$\mathcal{F}_{D,B}(f)(y', y_{d+1}) = \int_{\mathbb{R}_+^{d+1}} f(x', x_{d+1})\Lambda(x, y)d\mu_{k,\beta}(x). \qquad (52)$$

Remark 3.3.1. The transform $\mathcal{F}_{D,B}$ can also be written in the form

$$\mathcal{F}_{D,B} = \mathcal{F}_D \circ \mathcal{F}_B^\beta,$$

where $\mathcal{F}_D$ is the Dunkl transform given by Eq.(26) and $\mathcal{F}_B^\beta$ the Fourier-Bessel transform defined by

$$\forall\, \lambda \in \mathbb{R},\ \mathcal{F}_B(g)(\lambda) = \int_0^{+\infty} g(t) j_\beta(\lambda t) t^{2\beta+1} dt,\ g \in C_{*,c}(\mathbb{R}).$$

(See [16]).

The transform $\mathcal{F}_{D,B}$ satisfies the following properties:

Proposition 3.3.1. *i) For f in $L^1_{k,\beta}(\mathbb{R}^{d+1}_+)$ we have*

$$||\mathcal{F}_{D,B}(f)||_{k,\beta,\infty} \leq ||f||_{k,\beta,1}. \tag{53}$$

ii) For f in $\mathcal{S}_*(\mathbb{R}^{d+1})$ we have

$$\mathcal{F}_{D,B}(f) = \mathcal{F}_o \circ {}^t\mathcal{R}_{k,\beta}(f), \tag{54}$$

where $\mathcal{F}_o$ is the transform defined by : $\forall\, y = (y', y_{d+1}) \in \mathbb{R}^d \times [0, +\infty[$,

$$\mathcal{F}_o(f)(y', y_{d+1}) = \int_{\mathbb{R}^{d+1}_+} f(x', x_{d+1}) e^{-i<y',x'>} \cos(x_{d+1} y_{d+1}) dx' dx_{d+1},\ f \in C_{*,c}(\mathbb{R}^{d+1}). \tag{55}$$

iii) For f in $\mathcal{S}_*(\mathbb{R}^{d+1})$ we have

$$\forall y \in \mathbb{R}^{d+1}_+, \quad \mathcal{F}_{D,B}(\triangle_{k,\beta} f)(y) = -||y||^2 \mathcal{F}_{D,B}(f)(y). \tag{56}$$

From the previous properties of the Dunkl transform $\mathcal{F}_D$ and those of the Fourier-Bessel transform $\mathcal{F}_B^\beta(f)$ given in [16], we deduce the following theorems.

Theorem 3.3.1. *i) The Dunkl-Bessel transform $\mathcal{F}_{D,B}$ is a topological isomorphism, from $\mathcal{S}_*(\mathbb{R}^{d+1})$ onto itself.*

ii) The inverse transform $\mathcal{F}^{-1}_{D,B}$ is given by

$$\forall\, y \in \mathbb{R}^{d+1}_+, \quad \mathcal{F}^{-1}_{D,B}(f)(y) = m_{k,\beta} \mathcal{F}_{D,B}(f)(-y), \quad f \in S_*(\mathbb{R}^{d+1}),$$

where

$$m_{k,\beta} = \frac{c_k^2}{4^{\gamma+\beta+\frac{d}{2}} (\Gamma(\beta+1))^2}. \tag{57}$$

Theorem 3.3.2. *For all f in $L^1_{k,\beta}(\mathbb{R}^{d+1}_+)$ such that $\mathcal{F}_{D,B}(f)$ is in $L^1_{k,\beta}(\mathbb{R}^{d+1}_+)$, we have the inversion formula*

$$f(y) = m_{k,\beta} \int_{\mathbb{R}^{d+1}_+} \mathcal{F}_{D,B}(f)(x) \Lambda(x, -y) d\mu_{k,\beta}(x), \quad a.e. \tag{58}$$

Theorem 3.3.3. *i) (Plancherel formula for $\mathcal{F}_{D,B}$) For all f in $\mathcal{S}_*(\mathbb{R}^{d+1})$ we have*

$$\int_{\mathbb{R}_+^{d+1}} |f(x)|^2 d\mu_{k,\beta}(x) = m_{k,\beta} \int_{\mathbb{R}_+^{d+1}} |\mathcal{F}_{D,B}(f)(y)|^2 d\mu_{k,\beta}(y). \tag{59}$$

ii) (Plancherel theorem for $\mathcal{F}_{D,B}$) The renormalized Dunkl-Bessel transform $f \to m_{k,\beta}^{\frac{1}{2}}\mathcal{F}_{D,B}(f)$ can be uniquely extended to an isometric isomorphism on $L^2_{k,\beta}(\mathbb{R}_+^{d+1})$.

Proposition 3.3.2. *For all $f \in L^1_{k,\beta}(\mathbb{R}_+^{d+1})$, we have*

$$\mathcal{F}_{D,B}(f)(y) = \mathcal{F}_o \circ {}^t\mathcal{R}_{k,\beta}(f)(y), \quad y \in \mathbb{R}_+^{d+1}, \tag{60}$$

where $\mathcal{F}_o$ is the transform given by Eq.(55).

Proof

We obtain the result by applying Eq.(51) to the function $g(x) = e^{-i\langle x',y'\rangle}\cos(x_{d+1}y_{d+1})$ and by using the relation (42).

3.4. The Dunkl-Bessel harmonic polynomials.

We denote by P_m^{d+1} the set of homogeneous polynomials on $\mathbb{R}^{d+1}$ of degree m.

We say that a polynomial p in P_m^{d+1} is Dunkl-Bessel harmonic (D-B harmonic), if it satisfies $\triangle_{k,\beta}p = 0$.

We denote by $\mathcal{H}_m^{D,B}$ the set of polynomials in P_m^{d+1} which are D-B harmonic.

Theorem 3.4.1.*Each polynomial $\psi \in P_m^{d+1}$ has the unique representation*

$$\psi(x) = \sum_{s=0}^{[\frac{m}{2}]} (\frac{||x||}{2})^{2s} \widetilde{Z}_{m-2s,k,\beta}(x), \tag{61}$$

where $\widetilde{Z}_{m-2s,k,\beta}$ is a polynomial in $\mathcal{H}_{m-2s}^{D,B}$ given by

$$\widetilde{Z}_{m-2s,k,\beta}(x) = \sum_{j=0}^{[\frac{m}{2}]-s} a_{j,m,s}^{k,\beta} (\frac{||x||}{2})^{2j} \triangle_{k,\beta}^{s+j}\psi(x),$$

where

$$a_{j,m,s}^{k,\beta} = (-1)^j \frac{(m-2s+\gamma+\beta+\frac{d}{2})\Gamma(m-2s-j+\gamma+\beta+\frac{d}{2})}{s!j!\Gamma(m-s+\gamma+\beta+1+\frac{d}{2})} \tag{62}$$

(See [14].)

Theorem 3.4.2. *Let H be in $\mathcal{H}_m^{D,B}$. We have*

$$\mathcal{F}_{D,B}(e^{-\frac{||x||^2}{2}}H)(y) = \frac{i^m 2^{\gamma+\beta+\frac{d}{2}}\Gamma(\beta+1)}{c_k}e^{-\frac{||y||^2}{2}}H(y). \tag{63}$$

(See [13].)

Corollary 3.4.1. *Let $\psi \in P_m^{d+1}$, for all $\delta > 0$, there exists a polynomial Q on $I\!R^{d+1}$ such that*

$$\mathcal{F}_{D,B}(\psi\, e^{-\delta||x||^2})(y) = Q(y)e^{-\frac{1}{4\delta}||y||^2}.$$

Proof

From Theorem 3.4.1 and the relation (56) we have

$$\begin{aligned}
\forall\, y \in I\!R_+^{d+1},\ \mathcal{F}_{D,B}(\psi\, e^{-\delta||x||^2})(y) &= \mathcal{F}_{D,B}(\sum_{s=0}^{[\frac{m}{2}]}(\frac{||x||}{2})^{2s}\widetilde{Z}_{m-2s,k,\beta}(x)e^{-\delta||x||^2})(y),\\
&= \sum_{s=0}^{[\frac{m}{2}]}2^{-2s}\mathcal{F}_{D,B}(||x||^{2s}\widetilde{Z}_{m-2s,k,\beta}(x)e^{-\delta||x||^2})(y),\\
&= \sum_{s=0}^{[\frac{m}{2}]}2^{-2s}\triangle_{k,\beta}^{s}(\mathcal{F}_{D,B}(\widetilde{Z}_{m-2s,k,\beta}(x)e^{-\delta||x||^2}))(y).
\end{aligned}$$

On the other hand by applying Theorem 3.4.2 we obtain

$$\forall\ y \in I\!R_+^{d+1}, \quad \mathcal{F}_{D,B}(\widetilde{Z}_{m-2s,k,\beta}(x)e^{-\delta||x||^2})(y) = \frac{i^{m-2s}2^{2s-m-1}\Gamma(\beta+1)}{c_k\delta^{\gamma+\beta+\frac{d}{2}+m+1-2s}}\widetilde{Z}_{m-2s,k,\beta}(y)e^{-\frac{||y||^2}{4\delta}}.$$

By using the product rule (7), we deduce then there exists a polynomial Q on $I\!R_+^{d+1}$ such that

$$\mathcal{F}_{D,B}(\psi\, e^{-\delta||x||^2})(y) = Q(y)e^{-\frac{1}{4\delta}||y||^2}.$$

4. AN L^p VERSION OF HARDY'S THEOREM ASSOCIATED WITH THE DUNKL-BESSEL TRANSFORM

To prove the main result of this section we need the following lemmas of complex variables. The two first can be proved as in [7] and the third as Lemma 2.1 of [15] .

Lemma 4.1. *Let h be an entire function on $\mathbb{C}^{d+1}$ even with respect to the last variable such that*

$$\forall z \in \mathbb{C}^{d+1},\ |h(z)| \leq C\prod_{j=1}^{d+1}e^{a(Rez_j)^2},$$

and

$$\forall x \in I\!R_+^{d+1},\ |h(x)| \leq C,$$

for some $a > 0$ *and* $C > 0$*. Then* h *is constant on* $\mathbb{C}^{d+1}$.

Lemma 4.2. *Let* $p \in [1, +\infty[$ *and* h *an entire function on* $\mathbb{C}^{d+1}$ *even with respect to the last variable. We assume that:*
i) There exists $j \in \{1, ..., d+1\}$ *such that*

$$\forall z \in \mathbb{C}^{d+1},\ |h(z)| \leq M(z_1, ..., z_{j-1}, z_{j+1}, ..., z_{d+1}) e^{a(Rez_j)^2}, \tag{64}$$

for some $a > 0$ *and* M *a positive function on* $\mathbb{C}^{d+1}$.
ii) $||h_{/\mathbb{R}_+^{d+1}}||_{k,\beta,q} < +\infty$.
Then $h \equiv 0$.

Lemma 4.3. *Let* h *be an entire function on* $\mathbb{C}^{d+1}$ *even with respect to the last variable such that*

$$\forall z \in \mathbb{C}^{d+1}, \quad |h(z)| \leq Ce^{a||z||^2}, \tag{65}$$

and

$$\forall x \in \mathbb{R}^{d+1}, \quad |h(x)| \leq Ce^{-a||x||^2}, \tag{66}$$

for some positive constants a *and* C*. Then*

$$h(z) = Const.e^{-a(z_1^2 + ... + z_{d+1}^2)}, \quad z = (z_1, ..., z_{d+1}) \in C^{d+1}.$$

The following propositions are also necessary to obtain the result of this section.

Proposition 4.1. *Let* $a > 0$*. For all* $y \in \mathbb{R}_+^{d+1}$*, we have*

$$^t\mathcal{R}_{k,\beta}(e^{-a||x||^2})(y) = C(a)e^{-a||y||^2}, \tag{67}$$

where

$$C(a) = \frac{\Gamma(\beta+1)}{c_k a^{\gamma+\beta+\frac{1}{2}} \pi^{\frac{d+1}{2}}},$$

with c_k *the constant given by Eq.(5).*

Proof

As the function $e^{-a||x||^2}$ is in $\mathcal{S}_*(\mathbb{R}_+^{d+1})$, then from the relation (54) we show that

$$^t\mathcal{R}_{k,\beta}(e^{-a||x||^2})(y) = \mathcal{F}_o^{-1} \circ \mathcal{F}_{D,B}(e^{-a||x||^2})(y). \tag{68}$$

But from [13], p.460, we have

$$\mathcal{F}_{D,B}(e^{-a||x||^2})(y) = \frac{\Gamma(\beta+1)}{2c_k a^{\gamma+\beta+\frac{d}{2}+1}} e^{\frac{-||y||^2}{4a}}. \tag{69}$$

On the other hand since as for a regular function f on $I\!R_{+}^{d+1}$ we have

$$\mathcal{F}_o^{-1}(f)(y)=\frac{2^2}{(2\pi)^{d+1}}\mathcal{F}_o(f)(-y),$$

then we obtain Eq.(67) by applying to the relation (69) the inverse of the transform $\mathcal{F}_o$.

Proposition 4.2. *Let $p\in[1,+\infty]$ and f a measurable function on $I\!R_{+}^{d+1}$ such that $||e^{a||x||^2}f||_{k,\beta,p}<+\infty$ for some $a>0$. Then*

$$||e^{a||x||^2}\ {}^t\mathcal{R}_{k,\beta}(f)||_p<+\infty,$$

where $||.||_p$ is the norm of the usual Lebesgue space $L^p(I\!R_{+}^{d+1})$.

Proof

From the hypothesis it follows that $f\in L^1_{k,\beta}(I\!R_{+}^{d+1})$. Then by Theorem 3.2.1, the function ${}^t\mathcal{R}_{k,\beta}(f)$ is defined almost everywhere on $I\!R_{+}^{d+1}$. Now we distinguish two cases:

i) If $p\in[1,+\infty[$, we have

$$||e^{a||x||^2}\ {}^t\mathcal{R}_{k,\beta}(f)||_p^p\leq\int_{I\!R_{+}^{d+1}}e^{ap||x||^2}\Big(\int_{I\!R_{+}^{d+1}}f(y)d\varrho_x^{k,\beta}(y)\Big)^p dx.$$

On the other hand by using the Hölder's inequality we have

$$\begin{aligned}\Big|\int_{I\!R_{+}^{d+1}}f(y)d\varrho_x^{k,\beta}(y)\Big|^p &\leq \Big(\int_{I\!R_{+}^{d+1}}e^{ap||y||^2}|f(y)|^p d\varrho_x^{k,\beta}(y)\Big)\Big(\int_{I\!R_{+}^{d+1}}e^{-ap'||y||^2}d\varrho_x^{k,\beta}(y)\Big)^{\frac{p}{p'}}\\ &\leq \big({}^t\mathcal{R}_{k,\beta}(e^{ap||y||^2}|f|^p)(x)\big)\big({}^t\mathcal{R}_{k,\beta}(e^{-ap'||y||^2})(x)\big)^{\frac{p}{p'}},\end{aligned}$$

where p' is the conjugate exponent of p.

Thus from Proposition 4.1 and the relation (67) we obtain

$$||e^{a||x||^2}\ {}^t\mathcal{R}_{k,\beta}(f)||_p^p\leq(C(ap'))^{\frac{p}{p'}}\int_{I\!R_{+}^{d+1}}{}^t\mathcal{R}_{k,\beta}(e^{ap||y||^2}|f|^p)(x)dx.$$

But from Eq.(51) we have

$$\int_{I\!R_{+}^{d+1}}{}^t\mathcal{R}_{k,\beta}(e^{ap||y||^2}|f|^p)(x)dx=\int_{I\!R_{+}^{d+1}}e^{ap||y||^2}|f|^p(y)d\mu_{k,\beta}(y),\tag{70}$$

then

$$||e^{a||x||^2}\ {}^t\mathcal{R}_{k,\beta}(f)||_p\leq(C(ap'))^{\frac{1}{p'}}||e^{a||y||^2}f||_{k,\beta,p}<+\infty.$$

ii) If $p=+\infty$, for almost all $x\in I\!R_{+}^{d+1}$ we have

$$\begin{aligned}|{}^t\mathcal{R}_{k,\beta}(f)(x)| &\leq \Big(\int_{I\!R_{+}^{d+1}}e^{a||y||^2}|f(y)|e^{-a||y||^2}d\varrho_x^{k,\beta}(y)\Big)\\ &\leq {}^t\mathcal{R}_{k,\beta}(e^{-a||y||^2})(x)||e^{a||y||^2}f||_{k,\beta,\infty},\end{aligned}$$

thus from Proposition 4.1, we obtain

$$e^{a||x||^2}|^t\mathcal{R}_{k,\beta}(f)(x)| \leq C(a)||e^{a||y||^2}f||_{k,\beta,\infty} < +\infty,$$

where $C(a)$ is the constant of Eq.(67). This completes the proof.

Proposition 4.3. *Let $p \in [1,+\infty]$ and f a measurable function on $\mathbb{R}^{d+1}_+$ such that $||e^{a||x||^2}f||_{k,\beta,p} < +\infty$ for some $a > 0$. Then the function defined on $\mathbb{C}^{d+1}$ by*

$$\mathcal{F}_{D,B}(f)(z) = \int_{\mathbb{R}^{d+1}_+} f(x)\Lambda(x,z)d\mu_{k,\beta}(x) \tag{71}$$

is entire on $\mathbb{C}^{d+1}$ and even with respect to the last variable. Moreover there exists a positive constant C such that for all $\xi,\eta \in \mathbb{R}^{d+1}_+$, we have

$$|\mathcal{F}_{D,B}(\xi+i\eta)| \leq Ce^{\frac{||\eta||^2}{4a}}. \tag{72}$$

Proof

From the derivation theorem under the integral sign and Hölder's inequality we deduce that the function defined on $\mathbb{C}^{d+1}$ by Eq.(71) is entire on $\mathbb{C}^{d+1}$ and even with respect to the last variable.

On the other hand, since $f \in L^1_{k,\beta}(\mathbb{R}^{d+1}_+)$, then from Eq.(60) we deduce that for all $\xi,\eta \in \mathbb{R}^{d+1}_+$, we have

$$\begin{aligned}
\mathcal{F}_{D,B}(\xi+i\eta) &= \mathcal{F}_o \circ {}^t\mathcal{R}_{k,\beta}(f)(\xi+i\eta), \\
&= \int_{\mathbb{R}^{d+1}_+} {}^t\mathcal{R}_{k,\beta}(f)(x)e^{-i\langle x',\xi'+i\eta'\rangle}\cos(x_{d+1}(\xi_{d+1}+i\eta_{d+1}))dx'dx_{d+1}. \\
&= \frac{1}{2}\int_{\mathbb{R}^{d+1}_+} {}^t\mathcal{R}_{k,\beta}(f)(x)(e^{-i\langle x,\xi\rangle}e^{\langle x,\eta\rangle} + e^{-i\langle x,\widetilde{\xi}\rangle}e^{\langle x,\widetilde{\eta}\rangle})dx'dx_{d+1},
\end{aligned}$$

where $\widetilde{\eta} = (\eta', -\eta_{d+1})$ and $\widetilde{\xi} = (\xi', -\xi_{d+1})$.

Thus

$$|\mathcal{F}_{D,B}(\xi+i\eta)| \leq \frac{1}{2}e^{\frac{||\eta||^2}{4a}}\int_{\mathbb{R}^{d+1}_+} e^{a||x||^2}\,|^t\mathcal{R}_{k,\beta}(f)(x)|(e^{-a||x-\frac{\eta}{2a}||^2} + e^{-a||x-\frac{\widetilde{\eta}}{2a}||^2})dx.$$

Using Hölder's inequality and Proposition 4.2, we obtain

$$|\mathcal{F}_{D,B}(\xi+i\eta)| \leq \frac{1}{2}e^{\frac{||\eta||^2}{4a}}||e^{a||x||^2}\,{}^t\mathcal{R}_{k,\beta}(f)||_p(\int_{\mathbb{R}^{d+1}_+}(e^{-ap'||x-\frac{\eta}{2a}||^2} + e^{-ap'||x-\frac{\widetilde{\eta}}{2a}||^2}dx)^{\frac{1}{p'}},$$

where p' is the conjugate exponent of p. We deduce the result from this inequality.

Theorem 4.1. *Let f be a measurable function on $\mathbb{R}^{d+1}_+$ such that*

$$||e^{a||x||^2}f||_{k,\beta,p} < +\infty \quad and \quad ||e^{b||y||^2}\mathcal{F}_{D,B}(f)||_{k,\beta,p} < +\infty, \tag{73}$$

for some constants $a>0, b>0$, and $1\leq p,q\leq +\infty$ at least one of them is finite.

Then
i) If $ab\geq \frac{1}{4}$, we have $f=0$ a.e.
ii) If $ab<\frac{1}{4}$, for all $\delta\in]a,\frac{1}{4b}[$, the functions $f(x)=P(x)e^{-\delta||x||^2}$, where P is an arbitrary polynomial on $I\!R^{d+1}_+$, satisfy Eq.(73).

Proof

We shall divide the proof in several steps.

1) **First step:** $ab>\frac{1}{4}$

Consider the function h defined on $\mathbb{C}^{d+1}$ by

$$h(z)=\prod_{j=1}^{d+1} e^{\frac{z_j^2}{4a}}\mathcal{F}_{D,B}(f)(z). \tag{74}$$

This function is entire on $\mathbb{C}^{d+1}$ even with respect to the last variable and using Eq.(72) we obtain

$$\forall\, \xi,\, \eta\in I\!R^{d+1}_+,\ \ |h(\xi+i\eta)|\leq Ce^{\frac{||\xi||^2}{4a}}. \tag{75}$$

In the following we consider two cases:

i) If $q<+\infty$, we have

$$\begin{aligned} ||h_{/I\!R^{d+1}_+}||^q_{k,\beta,q} &= \int_{I\!R^{d+1}_+} |e^{\frac{||y||^2}{4a}}\mathcal{F}_{D,B}(f)(y)|^q d\mu_{k,\beta}(y)\\ &= \int_{I\!R^{d+1}_+} |e^{b||y||^2}\mathcal{F}_{D,B}(f)(y)|^q e^{q(\frac{1}{4a}-b)||y||^2} d\mu_{k,\beta}(y). \end{aligned}$$

Using the fact that $ab>\frac{1}{4}$ and the hypothesis (73), we obtain

$$||h_{/I\!R^{d+1}_+}||_{k,\beta,q}\leq ||e^{b||y||^2}\mathcal{F}_{D,B}(f)(y)||_{k,\beta,q}<+\infty. \tag{76}$$

From the relations (75) and (76), and Lemma 4.2 it follows that $h(z)=0$ for all $z\in\mathbb{C}^{d+1}$. Thus $\mathcal{F}_{D,B}(f)(y)=0$ for all $y\in I\!R^{d+1}_+$. The injectivity of $\mathcal{F}_{D,B}$ implies the result of the theorem in this case.

ii) If $q=+\infty$, we have

$$\begin{aligned} ||h_{/I\!R^{d+1}_+}||_{k,\beta,\infty} &= ess\ sup_{y\in I\!R^{d+1}_+}|e^{\frac{||y||^2}{4a}}\mathcal{F}_{D,B}(f)(y)|\\ &= ess\ sup_{y\in I\!R^{d+1}_+}\{|e^{b||y||^2}\mathcal{F}_{D,B}(f)(y)|e^{(\frac{1}{4a}-b)||y||^2}\}. \end{aligned}$$

Since $ab>\frac{1}{4}$, then from Eq.(73):

$$||h_{/I\!R^{d+1}_+}||_{k,\beta,\infty}\leq ||e^{b||y||^2}\mathcal{F}_{D,B}(f)(y)||_{k,\beta,\infty}<+\infty. \tag{77}$$

From Eqs.(75), (77) and Lemma 4.1, there exists a positive constant C such that for all $y \in \mathbb{R}_+^{d+1}$, $h(y) = C$. On the other hand, from Eq.(74) we have

$$\forall y \in \mathbb{R}_+^{d+1}, \ \mathcal{F}_{D,B}(f)(y) = Ce^{\frac{-||y||^2}{4a}}. \tag{78}$$

But the assumption on $\mathcal{F}_{D,B}(f)$ is expressed as

$$|\mathcal{F}_{D,B}(f)(y)| \leq Me^{-b||y||^2} \ a.e, \tag{79}$$

for some constant $M > 0$. The continuity of $\mathcal{F}_{D,B}(f)$ on $\mathbb{R}_+^{d+1}$ shows that the inequality (79) holds every where. By Eqs.(78) and (79) this is impossible since $ab > \frac{1}{4}$, unless if $C = 0$. Thus $\mathcal{F}_{D,B}(f)(y) = 0$ everywhere and then $f = 0$ a.e. on $\mathbb{R}_+^{d+1}$.

2) Second step $ab = \frac{1}{4}$.

i) If $1 \leq p \leq +\infty$ and $1 \leq q < +\infty$, the same proof as for the point i) of the first step implies that $f = 0$ a.e on $\mathbb{R}_+^{d+1}$.

ii) If $1 \leq p < +\infty$ and $q = +\infty$, from Proposition 4.2, Proposition 3.3.2 and Eq.(73) we deduce that the function ${}^t\mathcal{R}_{k,\beta}$ verifies

$$||e^{a||x||^2} \ {}^t\mathcal{R}_{k,\beta}(f)||_p < +\infty \quad \text{and } ||e^{b||y||^2} \ \mathcal{F}_o({}^t\mathcal{R}_{k,\beta}(f))||_\infty < +\infty,$$

where $\mathcal{F}_o$ is the Fourier transform given by Eq.(55).

We remark that by using the same proof as in [6] we can prove Hardy's theorem for the transform $\mathcal{F}_o$. Using this result, we deduce that ${}^t\mathcal{R}_{k,\beta}(f)(x) = 0$ a.e, on $\mathbb{R}_+^{d+1}$. Thus $\mathcal{F}_{D,B}(f)(y) = 0$ for all $y \in \mathbb{R}_+^{d+1}$, which implies that $f = 0$ a.e. on $\mathbb{R}_+^{d+1}$.

3)Third step: $ab < \frac{1}{4}$

Let P be a polynomial on $\mathbb{R}^{d+1}$ and $\delta \in]a, \frac{1}{4b}[$. From Corollary 3.4.1 we deduce that there exists a polynomial $\widetilde{Q}$ on $\mathbb{R}^{d+1}$ such that

$$\forall \ y \in \mathbb{R}_+^{d+1}, \quad \mathcal{F}_{D,B}(\mathrm{P}e^{-\delta||x||^2})(y) = \widetilde{Q}(y)e^{-\frac{1}{4\delta}||y||^2}.$$

It is clear that the function $f(x) = \mathrm{P}(x)e^{-\delta||x||^2}$ satisfies the conditions (73). This completes the proof of Theorem 4.1.

We determine now the functions f satisfying Eq.(73) in the special case $p = q = +\infty$.

Theorem 4.2. *Let f be a measurable function on $\mathbb{R}_+^{d+1}$ such that*

$$|f(x)| \leq Me^{-a||x||^2} a.e \text{ and} \, \forall y \in \mathbb{R}_+^{d+1}, \ |\mathcal{F}_{D,B}(f)(y)| \leq Me^{-b||y||^2}, \tag{80}$$

for some constants $a > 0$, $b > 0$ and $M > 0$. Then

i) If $ab > \frac{1}{4}$, we have $f = 0$ a.e.

ii) If $ab = \frac{1}{4}$, the function f is of the form $f(x) = C_0 e^{-a||x||^2}$, for some real constant C_0.

iii) If $ab < \frac{1}{4}$, there are infinity many nonzero functions f satisfying the conditions (80).

Proof

i) If $ab > \frac{1}{4}$, the point ii) of the first step of the proof of Theorem 4.1 gives the result.

ii) From Eq.(80), Proposition 4.1 and Proposition 3.3.2 , the function ${}^t\mathcal{R}_{k,\beta}(f)$ satisfies

$$|{}^t\mathcal{R}_{k,\beta}(f)(x)| \leq C(a)Me^{-a||x||^2} \text{a.e and } \forall\, y\; \in I\!R^{d+1}_+,\; |\mathcal{F}_o({}^t\mathcal{R}_{k,\beta}(f))(y)| \leq Me^{-b||y||^2},$$

where $C(a)$ is the constant in the formula (67). Using Hardy's theorem for the transform $\mathcal{F}_o$ which can be deduced from Lemma 4.3. Then

$${}^t\mathcal{R}_{k,\beta}(f)(x) = C_1 e^{-a||x||^2},$$

where C_1 is a real constant. We deduce from (60) that there exists $C_2 \in I\!R$ such that

$$\mathcal{F}_{D,B}(f)(y) = C_2 e^{-\frac{1}{4a}||y||^2}.$$

Thus by using formula (69) we have $f(x) = C_0 e^{-a||x||^2}$, with C_0 is a real constant. The result of the point ii) is proved.

iii) If $ab < \frac{1}{4}$, the functions defined in the third step of the proof of Theorem 4.1 clearly satisfies also the conditions (80). This completes the proof of Theorem 4.2.

Acknowledgements

We would like to think Professor Dr.Virginia Kiryakova of Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, for her interesting remarks and her help and attention.

References

[1] **M.G.Cowling and J.F.Price,** *Generalization of Heisemberg inequality, Lecture notes in Math. 992. Springer, Berlin, 1983, pp.443-449.*

[2] **M.F.E.de Jeu,** *The Dunkl transform, Invent.Math. 113 , 147-162, (1993).*

[3] **C. F. Dunkl,** *Reflection groups and orthogonal polynomials on the sphere, Math. Z. 197, 33-60, (1988).*

[4] **C. F. Dunkl,** *Differential-difference operators associated to reflection group, Trans. Am. Math. Soc. 311, 167 - 183, (1989).*

[5] **C.F.Dunkl,** *Hankel transforms associated to finite reflection groups, Contemp. Math. 138, 123-138, (1992).*

[6] **M.Eguchi, S.Koizumi and K.Kumahara,** *An L^p version of the Hardy theorem for the motion group, J.Austral.Math.Soc.(Series A) 68, 55-67, (2000).*

[7] **L.Gallardo and K.Trimèche,** *Un analogue d'un théorème de Hardy pour la transformation de Dunkl, C.R.Acad.Sci.Paris, t.334, Serie I, 849-854, (2002).*

[8] **G.H.Hardy,** *A theorem concerning Fourier transform, J.London Math Soc. 8, 227-231, (1933).*

[9] **G.J.Heckman,** *An elementary approach to the hypergeometric shift operators of Opdam, Invent.Math. 103, 341-350, (1991).*

[10] **K.Hikami,** *Dunkl operator formalism for quantum many-body problems associated with classical root systems,J.Phys.SoS.Japan, 65, 394-401, (1996).*

[11] **S.Kakei,***Common algebraic structure for the Calogera-Sutherland models, J.Phys. A 29, 619-624, (1996).*

[12] **L. Lapointe and L.Vinet,***Exact operator solution of the Calogera-Sutherland model, Comm.Math.Phys.178, 425-452, (1996).*

[13] **H. Mejjaoli and K. Trimèche,** *Harmonic analysis associated with the Dunkl-Bessel-Laplace operator and a mean value property, Fract.Calc.Appl.Anal. Vol 4, (4), 443-480, (2001).*

[14] **H. Mejjaoli and K. Trimèche,** *Spherical harmonics associated with the Dunkl-Bessel operator, Preprint of the Faculty of Sciences of Tunis, 2002.*

[15] **A.Sitaram and M.Sundari,** *An analogue of Hardy's theorem for very rapidly decreasing functions on semi-simple Lie groups, Pacific J. Math.177, 178-200, (1997).*

[16] **K. Trimèche,***Generalized Harmonic Analysis and Wavelet Packets, Gordon and Breach Science Publishars, 2001.*

[17] **K. Trimèche,***The Dunkl intertwining operator on spaces of functions and distributions and integral representation of its dual, Integ. Transf. and Special Funct. Vol. 12, (4), 349-374, (2001).*

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.4,419-427,2004,COPYRIGHT 2004 EUDOXUS PRESS,LLC

Fractional Green's Function for the Fractional Differential Equation with Constant Coefficients

Zaid Odibat
Faculty of Applied Science
Al-Balqa Applied University
Salt – Jordan

Tel: +962 5 3532586 Ext. 3728
Fax: +962 5 3557349
e-mail: *odibat@bau.edu.jo*

Ahmed Alawneh
Dept. of Mathematics
University of Jordan
Amman – Jordan

Tel: +962 6 5355000 Ext. 2622
Fax: +962 6 5348932
e-mail: *alawneh@sci.ju.edu.jo*

Abstract

We consider the fractional differential equation, that is obtained from the nth order ordinary differential equation with constant coefficients by replacing the ordinary derivative D by the fractional derivative D^{α}, $0<\alpha<1$ (in the Riemman-Liouville sense). We give the relationship between the fractional Green's function for the fractional differential equation and the Green's function for the ordinary differential equation. Some applications are also presented.

Key words. fractional differential equation, fractional Green's function, H-function, generalized Mittag-Leffler function

1. Introduction

The Green's function $G(t)$ related to the ordinary differential equation with constant coefficient

$$\left[\, D^n + a_1 D^{n-1} + \cdots + a_n D^0 \,\right] \Phi(t) = f(t)\,, \tag{1}$$

where a_i's are arbitrary constants, with the initial conditions

$$\Phi(0) = D\,\Phi(0) = \cdots = D^{n-1}\,\Phi(0) = 0\,, \tag{2}$$

is found by the inverse Laplace transform of the following expression

$$\widetilde{G}(p) = \frac{1}{Q(p)}\,, \tag{3}$$

where

$$Q(x) = x^n + a_1\, x^{n-1} + \cdots + a_0\,. \tag{4}$$

Miller and Ross [9] have proved that the fractional Green's function $G_\alpha(t)$ related to the fractional differential equation

$$\left[\, D^{n\alpha} + a_1 D^{(n-1)\alpha} + \cdots + a_n D^0 \,\right] \Phi(t) = f(t)\;, \qquad t > 0 \tag{5}$$

where D^α is the Riemman-Liouville fractional derivative of order $0 < \alpha < 1$, with the initial conditions

$$\Phi(0) = D\,\Phi(0) = \cdots = D^{N-1}\,\Phi(0) = 0\,, \tag{6}$$

where N is the smallest positive integer greater or equal $n\alpha$, is found by the inverse Laplace transform of the following expression

$$\widetilde{G}_\alpha(p) = \frac{1}{Q(p^\alpha)}\,. \tag{7}$$

Also, Miller and Ross have proved that

$$\Phi(t) = \int_0^t G_\alpha(t-\tau)\; f(\tau)\, d\tau \tag{8}$$

is the unique solution for equation (5).

Relations (3) and (7), shows that Laplace transform $\widetilde{G}_\alpha(p)$ of the Green's function $G_\alpha(t)$ for the fractional differential equation (5) and Laplace transform $\widetilde{G}(p)$ of the Green's function $G(t)$ for the ordinary differential equation (1) are related by

$$\widetilde{G}_\alpha(p) = \widetilde{G}(p^\alpha)\,. \tag{9}$$

The Laplace and Mellin transforms of a function Φ on $\boldsymbol{R}^+$ are defined by

$$\tilde{\Phi}(p) = \int_0^\infty e^{-pt}\,\Phi(t)\,dt\,, \quad \operatorname{Re}(p) > 0 \tag{10}$$

and

$$\hat{\Phi}(s) = \int_0^\infty t^{s-1}\,\Phi(t)\,dt\,, \quad \operatorname{Re}(s) > 0 \tag{11}$$

and they are related by

$$\hat{\Phi}(s) = \frac{1}{\Gamma(1-s)} \int_0^\infty p^{-s}\,\Phi(p)\,dp\,. \tag{12}$$

Our aim here is to obtain the relationship between the fractional Green's function related to the fractional differential equation (5) and the Green's function related to the ordinary differential equation (1), then to use this relationship to find the fractional Green's function for some fractional differential equations.

2. Fractional Green's function

Theorem 1. The Mellin transforms $\hat{G}_\alpha(s)$ of the Green's function $G_\alpha(t)$ and $\hat{G}(s)$ of the Green's function $G(t)$ are related by

$$\hat{G}_\alpha(s) = \frac{1}{\alpha}\,\frac{\Gamma(1-(s/\alpha - 1/\alpha + 1))}{\Gamma(1-s)}\,\hat{G}(s/\alpha - 1/\alpha + 1)\,. \tag{13}$$

Proof: From (9) and (12), we get

$$\hat{G}_\alpha(s) = \frac{1}{\Gamma(1-s)} \int_0^\infty p^{-s}\,\tilde{G}(p^\alpha)\,dp\,. \tag{14}$$

The change of variables

$$q = p^\alpha \tag{15}$$

leads to

$$\hat{G}_\alpha(s) = \frac{1}{\alpha}\,\frac{1}{\Gamma(1-s)} \int_0^\infty q^{\frac{-s}{\alpha} + \frac{1}{\alpha} - 1}\,\tilde{G}(q)\,dq\,. \tag{16}$$

using (12), we have

$$\int_0^\infty q^{\frac{-s}{\alpha}+\frac{1}{\alpha}-1}\,\tilde{G}(q)\,dq=\Gamma\big(1-(s/\alpha-1/\alpha+1)\big)\;\hat{G}(s/\alpha-1/\alpha+1)\,. \tag{17}$$

Substitute (17) into (16), we obtain (13).

Theorem 2. The Green's functions $G_\alpha(t)$ and $G(t)$ are related by

$$G_\alpha(t)=t^{-1}\int_0^\infty G(z)\,g_\alpha(t^{-\alpha}\,z)\,dz\,, \tag{18}$$

where

$$g_\alpha(z)=H_{11}^{10}\left(\,z\,\middle|\,\begin{matrix}(0,\alpha)\\(0,1)\end{matrix}\right),$$

(19)

H_{11}^{10} is the H-function (see Appendix A). For $0<\alpha<1$, $g_\alpha(t)$ is an entire function and it has the power series representation

$$g_\alpha(t)=\sum_{k=0}^{\infty}\frac{(-1)^k\,t^k}{\Gamma(-\alpha k)\;k!}\,. \tag{20}$$

Proof: From the definition of the H-function, the Mellin transform of (19) is given by

$$\hat{g}_\alpha(s)=\frac{\Gamma(s)}{\Gamma(\alpha s)}\,. \tag{21}$$

Then, using (21), (13) can be written as

$$\hat{G}_\alpha(s)=\frac{1}{\alpha}\,\hat{g}_\alpha\big(1-(s/\alpha-1/\alpha+1)\big)\;\hat{G}(x,s/\alpha-1/\alpha+1)\,. \tag{22}$$

Next, using the following properties of the Mellin transform

$$\mathrm{M}\left(t^{-a}\;f(t)\right)=\hat{f}(s-a)\,, \tag{23}$$

$$\mathrm{M}\left(f(t^{\gamma})\right)=\frac{1}{\gamma}\,\hat{f}\left(\frac{s}{\gamma}\right), \tag{24}$$

and

$$\mathrm{M}\left(\int_0^\infty f(t\,y)\,g(y)\,dy\right)=\hat{f}(s)\,\hat{g}(1-s)\,. \tag{25}$$

The inverse Mellin transform for (22) gives

$$G_\alpha(t) = t^{\alpha-1} \int_0^\infty G(t^\alpha\ y)\, g_\alpha(y)\, dy\,. \tag{26}$$

The change of variables

$$z = t^\alpha\ y \tag{27}$$

leads to (18).

3. Applications

1. *Consider the second order ordinary differential equation*

$$D^2\ \Phi(t) + (a+b)\cdot D\,\Phi(t) + ab\cdot\Phi(t) = f(t)\,, \qquad (a\,, b > 0) \tag{28}$$

and the initial conditions

$$\Phi\,(0) = D\,\Phi(0) = 0\,. \tag{29}$$

The Greens function for this problem is

$$G(t) = \begin{cases} \dfrac{1}{b-a}\left[e^{-a\,t} - e^{-b\,t}\right] & if \quad a \neq b \\ t\,e^{-a\,t} & if \quad a = b \end{cases}. \tag{30}$$

According to theorem 2 the fractional Green's function $G_\alpha(t)$ for the fractional differential equation

$$D^{2\alpha}\ \Phi(t) + (a+b)\cdot D^\alpha\ \Phi(t) + ab\cdot\Phi(t) = f(t)\,, \qquad (t > 0) \tag{31}$$

with the initial conditions (29), where $a\,, b > 0$ and $0 < \alpha < 1$, is given by

$$G_\alpha(t) = t^{-1}\int_0^\infty G(z)\, H_{11}^{10}\left(t^{-\alpha} z \,\middle|\, \begin{matrix} (0\,,\alpha) \\ (0\,,1) \end{matrix} \right) dz\,, \tag{32}$$

From the property (60), we have

$$\int_0^\infty e^{-p\,z}\, H_{11}^{10}\left(z \,\middle|\, \begin{matrix} (0\,,\alpha) \\ (0\,,1) \end{matrix} \right) dz = \frac{1}{p}\, H_{12}^{10}\left(p \,\middle|\, \begin{matrix} (1{,}1) \\ (1{,}1)(1{,}\alpha) \end{matrix} \right), \quad \mathrm{Re}\,(p) > 0 \tag{33}$$

and

$$\int_0^\infty p\,e^{-p\,z}\, H_{11}^{10}\left(z \,\middle|\, \begin{matrix} (0\,,\alpha) \\ (0\,,1) \end{matrix} \right) dz = \frac{d}{dp}\left[\ \frac{1}{p}\, H_{12}^{10}\left(p \,\middle|\, \begin{matrix} (1{,}1) \\ (1{,}1)(1{,}\alpha) \end{matrix} \right) \right], \qquad \mathrm{Re}\,(p) > 0$$

(34)

Now, using the series representation of H-function, (33) and (34) can be written as

$$\int_0^\infty e^{-pz}\, H_{11}^{10}\left(z \,\middle|\, \begin{matrix} (0,\alpha) \\ (0,1) \end{matrix} \right) dz = E_{\alpha,\alpha}(-p)\,, \qquad \mathrm{Re}\,(p) > 0 \tag{35}$$

and

$$\int_0^\infty p\, e^{-pz}\, H_{11}^{10}\left(z \,\middle|\, \begin{matrix} (0,\alpha) \\ (0,1) \end{matrix} \right) dz = \frac{d}{dp}\, E_{\alpha,\alpha}(-p)\,, \qquad \mathrm{Re}\,(p) > 0 \tag{36}$$

where $E_{\alpha,\beta}(z)$ is the Mittag-Leffler function in the two parameters [6], [10].

It follows from (35), (36) and the properties of Laplace transform that the final expression for the fractional Green's function (32) can be written as

$$G_\alpha(t) = \begin{cases} \dfrac{t^{\alpha-1}}{b-a}\left[E_{\alpha,\alpha}(-a\,t^\alpha) - E_{\alpha,\alpha}(-b\,t^\alpha) \right] & if \quad a \neq b \\ t^{\alpha-1}\, \dfrac{d}{dx} E_{\alpha,\alpha}(-x\,t^\alpha)\Big|_{x=a} & if \quad a = b \end{cases} . \tag{37}$$

2. *Consider the second order ordinary differential equation*

$$D^2\, \Phi(t) + 2\lambda \cdot D\, \Phi(t) + (\lambda^2 + \mu^2) \cdot \Phi(t) = f(t)\,, \quad (\lambda > 0\,, \mu \neq 0) \tag{38}$$

and the initial conditions

$$\Phi\,(0) = D\,\Phi(0) = 0\,. \tag{39}$$

The Greens function for this problem is

$$G(t) = \frac{1}{\mu}\, e^{-\lambda t}\, \sin(\mu\, t)\,. \tag{40}$$

According to theorem 2 the fractional Green's function $G_\alpha(t)$ for the fractional differential equation

$$D^{2\alpha}\, \Phi(t) + 2\lambda \cdot D^\alpha\, \Phi(t) + (\lambda^2 + \mu^2) \cdot \Phi(t) = f(t)\,, \qquad (t > 0) \tag{41}$$

with the initial conditions (39), where $\lambda > 0$, $\mu \neq 0$ and $0 < \alpha < 1$, is given by

$$G_\alpha(t) = t^{-1} \int_0^\infty G(z)\, H_{11}^{10}\left(t^{-\alpha} z \,\middle|\, \begin{matrix} (0,\alpha) \\ (0,1) \end{matrix} \right) dz\,, \tag{42}$$

$$= \frac{t^{-1}}{\mu} \int_0^\infty e^{-\lambda z}\, \sin(\mu\, z)\, H_{11}^{10}\left(t^{-\alpha} z \,\middle|\, \begin{matrix} (0,\alpha) \\ (0,1) \end{matrix} \right) dz\,, \tag{43}$$

$$= \frac{t^{-1}}{2i\,\mu} \int_0^\infty \left[e^{(-\lambda+i\mu)z} - e^{-(\lambda+i\mu)z} \right] H_{11}^{10}\left(t^{-\alpha} z \,\middle|\, \begin{matrix} (0,\alpha) \\ (0,1) \end{matrix} \right) dz\,. \tag{44}$$

From (35) and the properties of Laplace transform, the fractional Green's function $G_\alpha(t)$ for the fractional differential equation (41) can be written as

$$G_\alpha(t) = \frac{t^{\alpha-1}}{2i\,\mu}\left[E_{\alpha,\alpha}\left((-\lambda+i\mu)t^\alpha\right) - E_{\alpha,\alpha}\left(-(\lambda+i\mu)t^\alpha\right)\right]. \tag{45}$$

4. Appendix A

Fox’s H-function

$$H_{pq}^{mn}(z) = H_{pq}^{mn}\left(z \,\middle|\, \begin{matrix} (a_j,\alpha_j) & j=1,\cdots,p \\ (b_i,\beta_i) & i=1,\cdots,q \end{matrix} \right), \tag{46}$$

with

$0 \le n \le p$, $1 \le m \le q$, a_j arbitrary complex numbers, α_j positive numbers, is characterized by its Mellin transform

$$\hat{H}_{pq}^{mn}(s) = \frac{A(s)\,B(s)}{C(s)\,D(s)}, \tag{47}$$

where

$$A(s) = \prod_{j=1}^{m} \Gamma(b_j + \beta_j s), \tag{48}$$

$$B(s) = \prod_{j=1}^{n} \Gamma(1 - a_j - \alpha_j s), \tag{49}$$

$$C(s) = \prod_{j=m+1}^{q} \Gamma(1 - b_j - \beta_j s), \tag{50}$$

and

$$D(s) = \prod_{j=n+1}^{p} \Gamma(a_j + \alpha_j s), \tag{51}$$

with empty products set equal to unity.

The set of the poles of $A(s)$ and $B(s)$, that are

$$P(A) = \left\{ s = -(b_j + k)/\beta_j\,, \quad j = 1,\cdots,m\,; \quad k = 0,1,\cdots \right\}, \tag{52}$$

$$P(B) = \left\{ s = (1 + k - a_j)/\alpha_j\,, \quad j = 1,\cdots,n\,; \quad k = 0,1,\cdots \right\}, \tag{53}$$

respectively, are supposed to be disjoint.

If the poles of $A(s)$ and $B(s)$ are simple and if:

$$\mu = \sum_{j=1}^{q} \beta_j - \sum_{j=1}^{n} \alpha_j > 0, \tag{54}$$

then $H_{pq}^{mn}(z)$ is an analytic function for $z \neq 0$ and has the series representation

$$H_{pq}^{mn}(z)=\sum_{h=1}^{m}\sum_{k=0}^{\infty}\frac{\prod_{j=1,j\neq h}^{m}\Gamma(b_j-\beta_j s_{hk})\prod_{j=1}^{n}\Gamma(1-a_j+\alpha_j s_{hk})}{\prod_{j=m+1}^{q}\Gamma(1-b_j+\beta_j s_{hk})\prod_{j=n+1}^{p}\Gamma(a_j-\alpha_j s_{hk})}\cdot\frac{(-1)^k z^{s_{hk}}}{k!\,\beta_h}, \tag{55}$$

where

$$s_{hk}=\frac{b_h+k}{\beta_k}. \tag{56}$$

The following identities for the H-function are well-known:

$$H_{pq}^{mn}\left(z\left|\begin{matrix}(a_j,\alpha_j)\\(b_i,\beta_i)\end{matrix}\right.\right)=H_{qp}^{nm}\left(\frac{1}{z}\left|\begin{matrix}(1-b_i,\beta_i)\\(1-a_j,\alpha_j)\end{matrix}\right.\right), \tag{57}$$

$$H_{pq}^{mn}\left(z\left|\begin{matrix}(a_j,\alpha_j)\\(b_i,\beta_i)\end{matrix}\right.\right)=k\,H_{pq}^{mn}\left(z^k\left|\begin{matrix}(a_j,k\alpha_j)\\(b_i,k\beta_i)\end{matrix}\right.\right), \tag{58}$$

$$z^{\delta}\,H_{pq}^{mn}\left(z\left|\begin{matrix}(a_j,\alpha_j)\\(b_i,\beta_i)\end{matrix}\right.\right)=H_{pq}^{mn}\left(z\left|\begin{matrix}(a_j+\delta\alpha_j,\alpha_j)\\(b_i+\delta\beta_i,\beta_i)\end{matrix}\right.\right). \tag{59}$$

The Laplace transform of the H-function is given by

$$L\,H_{pq}^{mn}\left(z\left|\begin{matrix}(a_j,\alpha_j)\\(b_i,\beta_i)\end{matrix}\right.\right)(p)=\frac{1}{p}H_{q\,p+1}^{n+1\,m}\left(p\left|\begin{matrix}(1-b_i,\beta_i)\\(1,1)(1-a_j,\alpha_j)\end{matrix}\right.\right),\ \text{for } 0<\mu\leq 1 \tag{60}$$

$$L\,H_{pq}^{mn}\left(z\left|\begin{matrix}(a_j,\alpha_j)\\(b_i,\beta_i)\end{matrix}\right.\right)(p)=\frac{1}{p}H_{p+1\,q}^{m\,n+1}\left(\frac{1}{p}\left|\begin{matrix}(0,1)(a_j,\alpha_j)\\(b_i,\beta_i)\end{matrix}\right.\right),\quad \text{for } \mu>1. \tag{61}$$

In particular, we mention

$$E_{\alpha,\beta}(z)=H_{12}^{11}\left(z\left|\begin{matrix}(0,1)\\(0,1)\,(1-\beta,\alpha)\end{matrix}\right.\right), \tag{62}$$

with $z\geq 0$, $0<\alpha\leq 1$, $\beta>\alpha$.

Using the series representation (55), we obtain

$$E_{\alpha,\beta}(z)=\sum_{k=0}^{\infty}(-1)^k\frac{z^k}{\Gamma(\alpha z+\beta)}, \tag{63}$$

which is the generalized Mittage-Leffler function.

References

1. Ch. Fox, The G and H functions as symmetrical Fourier kernels, *Trans. Amer. Math. Soc.*, 98 (1961), pp. 395-429.

2. K. B. Oldham and J. Spanier, *The fractional calculus*, Academic Press, New York, 1974.

3. L. M. Campos, On the solution of some simple fractional differential equations, *Internat. J. Math. & Math. Sci.*, 13, No 3 (1990), pp. 481-496.

4. M. Wyss and W. Wyss, Evolution, its fractional extension and generalization, *Fract. Calc. Appl. Anal.*, 4, No 3 (2001), pp. 273-284.

5. Podlubny, *Fractional differential equations*, Academic Press, San Diego, CA, 1999.

6. R. Gorenflo and F. Mainardi, *Fractional calculus: Integral and differential equations of fractional order:* From, *Fractals and fractional calculus*, Carpinteri & Mainardi, New York, 1997.

7. R. Gorenflo, Yu. Luchko and F. Mainardi, Analytical properties and applications of the Wrigth function. *Fract. Calc. Appl. Anal.*, 2, No 4 (1999), pp. 383-414.

8. S. G. Samko, A. A. Kilbas, and O.I. Marichev, *Fractional integrals and derivatives: Theory and applications*, Gordon and Breach, London, 1993.

9. S. Miller and B. Ross, *An introduction to the fractional calculus and fractional differential eqautions*, John Wiley & Sons, USA, 1993.

10. W. R. Schneider, Completely monotone generalized Mittag-Leffler functions, *Exposition. Math.*, 14, No 1 (1996), pp. 3-16.

11. W. Wyss, The fractional Black-Sholes equation, *Fract. Calc. Appl. Anal.*, 3, No 1 (2000), pp. 51-61.

12. Y. Luchko and R. Gorenflo, Scale-inveriant solutions of a partial differential equation of fractional order, *Fract. Calc. Appl. Anal.*, 1, No 1 (1998), pp. 63-77.

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts hard copies in triplicate, and in English, should be submitted to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel.001. 901.678.3144
e-mail:ganastss@memphis.edu.

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.
Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.
2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (VISIT www.msci.memphis.edu/~ganastss/jcaam / to save a copy of the style file.)They should be carefully prepared in all respects. Submitted copies should be brightly printed (not dot-matrix), double spaced, in ten point type size, on one side high quality paper 8(1/2)x11 inch. Manuscripts should have generous margins on all sides and should not exceed 24 pages.
3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.
4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.
The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.
6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes. The main body of the paper should begin on page no. 1, if possible.
7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.
8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right,and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.
9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.
Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.
10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.
References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,
name of publication, volume number, inclusive pages, and year of publication.
Authors should follow these examples:

<u>Journal Article</u>

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

<u>Book</u>

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

<u>Contribution to a Book</u>

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.
12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set

off from the text. Use the appropriate superscript numeral for citation in the text.
13. After each revision is made please again submit three hard copies of the revised manuscript, including in the final one. And after a manuscript has been accepted for publication and with all revisions incorporated, manuscripts, including the TEX/LaTex source file and the PDF file, are to be submitted to the Editor's Office on a personal-computer disk, 3.5 inch size. Label the disk with clearly written identifying information and properly ship, such as:

Your name, title of article, kind of computer used, kind of software and version number, disk format and files names of article, as well as abbreviated journal name.

Package the disk in a disk mailer or protective cardboard. Make sure contents of disks are identical with the ones of final hard copies submitted!

Note: The Editor's Office cannot accept the disk without the accompanying matching hard copies of manuscript. No e-mail final submissions are allowed! The disk submission must be used.

14. Effective 1 Nov. 2005 the journal's page charges are $8.00 per PDF file page. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown in the Scope and Prices section.

No galleys will be sent and the contact author will receive one(1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

TABLE OF CONTENTS,JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.2,NO.4,2004